National Library
of Canada

Canadian Theses Service

Ottawa, Canada
K1A 0N4

Bibliothèque nationale
du Canada

Services des thèses canadiennes

CANADIAN THESES

THÈSES CANADIENNES

NOTICE

The quality of this microfiche is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Previously copyrighted materials (journal articles, published tests, etc.) are not filmed.

Reproduction in full or in part of this film is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30. Please read the authorization forms which accompany this thesis.

AVIS

La qualité de cette microfiche dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

Les documents qui font déjà l'objet d'un droit d'auteur (articles de revue, examens publiés, etc.) ne sont pas microfilmés.

La reproduction, même partielle, de ce microfilm est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30. Veuillez prendre connaissance des formules d'autorisation qui accompagnent cette thèse.

THIS DISSERTATION
HAS BEEN MICROFILMED
EXACTLY AS RECEIVED

LA THÈSE A ÉTÉ
MICROFILMÉE TELLE QUE
NOUS L'AVONS REÇUE

NL 339 (r. 86/01)

Canada

250

National Library
of Canada

Canadian Theses Division

Ottawa, Canada
K1A 0N4

Bibliothèque nationale
du Canada

Division des thèses canadiennes

0-315-26981-2

PERMISSION TO MICROFILM — AUTORISATION DE MICROFILMER

- Please print or type — Écrire en lettres moulées ou dactylographier

Full Name of Author — Nom complet de l'auteur

FREDRICK FRANCIS ULMER

Date of Birth — Date de naissance

October 1, 1950

Country of Birth — Lieu de naissance

Canada

Permanent Address — Résidence fixe

7603 75 Avenue
Edmonton, Alberta
T6C 0G2

Title of Thesis — Titre de la thèse

Melody and Synchrony in the Text of Conversation

University — Université

University of Alberta

Degree for which thesis was presented — Grade pour lequel cette thèse fut présentée

M. Ed.

Year this degree conferred — Année d'obtention de ce grade

1983

Name of Supervisor — Nom du directeur de thèse

DR CARL A. URION

Permission is hereby granted to the NATIONAL LIBRARY OF CANADA to microfilm this thesis and to lend or sell copies of the film.

The author reserves other publication rights, and neither the thesis nor extensive extracts from it may be printed or otherwise reproduced without the author's written permission.

L'autorisation est, par la présente, accordée à la BIBLIOTHÈQUE NATIONALE DU CANADA de microfilmer cette thèse et de prêter ou de vendre des exemplaires du film.

L'auteur se réserve les autres droits de publication; ni la thèse ni de longs extraits de celle-ci ne doivent être imprimés ou autrement reproduits sans l'autorisation écrite de l'auteur.

Date

October 17, 1983

Signature

Fredrick Ulmer

NL-91 (4/77)

THE UNIVERSITY OF ALBERTA

Melody and Synchrony in the Text of Conversation

by

© Fredrick Francis Ulmer

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES AND RESEARCH
IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE
OF Master of Education
IN
Anthropology and Education

Department of Educational Foundations

EDMONTON, ALBERTA
Fall, 1983

THE UNIVERSITY OF ALBERTA

RELEASE FORM

NAME OF AUTHOR Fredrick Francis Ulmer

TITLE OF THESIS Melody and Synchrony in the Text of Conversation

DEGREE FOR WHICH THESIS WAS PRESENTED Master of Education

YEAR THIS DEGREE GRANTED Fall, 1983

Permission is hereby granted to THE UNIVERSITY OF ALBERTA LIBRARY to reproduce single copies of this thesis and to lend or sell such copies for private, scholarly or scientific research purposes only.

The author reserves other publication rights, and neither the thesis nor extensive extracts from it may be printed or otherwise reproduced without the author's written permission.

(SIGNED) Fredrick Ulmer

PERMANENT ADDRESS:

7603 75 Avenue
Edmonton, Alberta
T6C 0G2

DATED October 17 1983

THE UNIVERSITY OF ALBERTA

FACULTY OF GRADUATE STUDIES AND RESEARCH

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research, for acceptance, a thesis entitled "Melody and Synchrony in the Text of Conversation," submitted by Fredrick Francis Ulmer in partial fulfilment of the requirements for the degree of Master of Education in Anthropology and Education.

Carl Urion
Supervisor

Regna Darnell

A D Fisher

Date. 15 August 1983

Abstract

This study is an examination of several models of human speech interaction. A model recently proposed by Michael W. Mair is central to the study. Mair's comprehensive model includes several claims about synchrony, or body movement and acoustic parameters that are found to be temporally coincident when examined in microanalysis. Much of Mair's argument rests upon the claim that this synchrony is perceptually salient to the interactants and is evidence of "supra-individual" rhythms which can be seen to make conversation cohere; can in fact be said to indicate mutual "awareness."

In this study, as in Mair's, synchrony is defined as simultaneous manifestation of sound and movement between interactants, as well as simultaneous body movement and sound production--simultaneous within one tenth of a second--within one individual involved in interaction.

This study involves the analysis of three segments of dyadic conversation, each segment of about ten seconds duration. Measures of fundamental frequency of voice were compared to head and hand movements of each individual in dyadic conversation. It was found, when such a comparison was made, that the coincidences between these "channels" or "modalities," was very great; and that there was considerable coincidence between the interactants.

Acknowledgements

An undertaking such as this requires the creativity and support of many, and although I take full responsibility for the final form and expression of what follows, I would like to take this opportunity to express my gratitude to those who have helped along the way. I owe a particular debt to Dr. Carl A. Urion for his advice, counsel, and support at both intellectual and emotional levels throughout this enterprise. He has given more than these pages can show. I would also like to thank Dr. Regna Darnell and Dr. A. D. Fisher for their critical reading of this thesis. Their comments were essential in bringing the document to its final form. Dr. Michael W. Mair provided timely criticism and advice. His work is important to me and I hope I have done him justice. The technical staff in the Department of Linguistics and in the Instructional Technology Centre made the study possible. They went out of their way to help me.

In addition, I would like to express my gratitude to my family, and to the friends who helped me through. Special attention goes to Jeff Bullard who has been there throughout and who helped me in formulating and completing the study. Special thanks also are due Sophia Matiasz, the Urion family, Cora Weber-Pillwax and Cynthia MacCormac for their support.

Last but not least I would extend my thanks to Fernando P. Whitecloud, typist extraordinaire. Without him it would not have been done nearly so well.

Table of Contents

Chapter Page

1. Introduction 1

  1.1 Background To The Study 1

  1.2 Mair's Model of the Melody of the Text 6

    1.2.1 Approaching Mair's Model 7

    1.2.2 Early Research Into Communicative Interaction 11

    1.2.3 Recent Interaction Research 13

    1.2.4 Mair's Model 14

  1.3 The Present Study 19

  1.4 Organisation of the Thesis 29

2. Current Studies in Discourse and Prosody 31

  2.1 Speech Act Theory 31

  2.2 Acoustic Phonetics 33

    2.2.1 Isochrony Hypothesis 34

    2.2.2 The Comb and Chain Models 36

  2.3 Studies of Time and Perception in Speech 44

    2.3.1 Pauses and Timing: Methodological and Theoretical Problems 45

      2.3.1.1 Pauses as cognitive process 45

      2.3.1.2 Pause as affective domain 48

      2.3.1.3 Pause and timing as social phenomenon 48

      2.3.1.4 Limitations of "pause" studies 51

    2.3.2 Time as Coherence Principle: The Synthetic Approach 54

      2.3.2.1 Erickson and timing 55

      2.3.2.2 Scollon and tempo 56

      2.3.2.3 Critique of Scollon 62

2.3.2.4 Chapple and entrainment 67

2.3.2.5 Implications of Chapple's Model ..75

2.3.2.6 Critique of Chapple 77

2.3.3 Synchrony 81

3. Summary of Mair's Model 82

3.1 Faces and Gestalt Transforms 82

3.2 Speech Parameters, Visions, and Plans 87

3.2.1 Fundamental Frequency Contours 90

3.2.2 Rhythmicity and Plans 91

3.2.3 Rhythmicity and the Single Central Patterning Program In the Brain 97

3.2.4 Vision 99

3.2.5 Supra-glottal Modification 100

3.2.6 Synchrony 103

3.3 The Shared Text and Melody 106

3.3.1 Introduction 106

3.3.2 State of Play 112

3.3.3 Topic 119

3.3.4 Projection 120

3.3.5 The Linearity of Awareness 122

3.3.6 Summary of Rules Governing Speech Melody 126

3.3.7 Immersion 127

3.3.8 Summary of Principles 129

3.4 Rapport and Control 130

3.5 Tropisms 135

3.6 Summary 137

4. Critique of Mair's Model 143

4.1 Introduction 143
4.2 Descriptive Adequacy 146
    4.2.1 The GT Model 149
    4.2.2 Immersion 150
    4.2.3 Projection 151
    4.2.4 Charm and Charisma 152
    4.2.5 Ostensible Versus Oblique Topic 153
    4.2.6 Limitations 157
    4.2.7 Rapport 158
    4.2.8 Projection 159
    4.2.9 Tropisms 160
    4.2.10 Control as Dominance 160
4.3 Mair's Model as a Natural System Model 162
    4.3.1 Autopoetic and Allopoetic Systems 167
    4.3.2 Communicative and Linguistic Domains 170
    4.3.3 The Observer as a Natural System 171
    4.3.4 Implication of the Theory of Observation of Natural Systems 173
    4.3.5 Descriptive Adequacy of the Theory 176
    4.3.6 Summary 178
5. Methodology 179
    5.1 Initial Trials 179
    5.2 Selection of Subjects 187
    5.3 Subjects 190
    5.4 Protocols for Analysis 193
6. Selected Fragments of Text 208
    6.1 Introduction 208

6.2 Example 1 214

- 6.2.1 Sound and Movement Clusters 217
- 6.2.2 Text of the Selected Fragment 218
- 6.2.3 Sound and Movement Clusters: Interpretation 236

6.3 Example 2 249

- 6.3.1 Text of the Selected Fragment 252
- 6.3.2 Sound and Movement Clusters--Interpretation 282

6.4 Example 3 300

- 6.4.1 Sound and Movement Clusters--Description 303
- 6.4.2 Text of the Selected Fragment 303
- 6.4.3 Sound and Movement Clusters--Interpretation 332

7. Conclusions 357

7.1 Introduction 357

7.2 The Present Study 359

- 7.2.1 Contributions of the Study 361
  - 7.2.1.1 Synchrony 361
  - 7.2.1.2 Substantiation of the observation of synchrony 362
  - 7.2.1.3 Single central patterning program in the brain 363
  - 7.2.1.4 The social brain 363
  - 7.2.1.5 Autopoiesis 364
  - 7.2.1.6 Incorporation of other models ...364
- 7.2.2 Limitations 365
- 7.2.3 Implications For Further Research 367

Bibliography 368

List of Figures

Figure Page

3.1 The Necker Cube....................................85

3.2 Mair's Recording Set-Up............................89

3.3 Diagrammatic Representation of Force Field.........115

3.4 Plan and Percept...................................117

5.1 The Group Processes Laboratory.....................188

5.2 Equipment Set-Up For Mingograph Record.............203

6.1 Musical Notation of "Sick place, man"..............241

6.2 Musical Notation of "What did I read"..............244

6.3 Musical Notation of "Sick place"...................248

6.4 Musical Notation of "And depending.."..............288

6.5 Musical Notation of "In Ed Psych [snap]"...........292

6.6 Musical Notation of "Here, as a..."................298

6.7 Musical Notation for "...little creeps"............340

6.8 Musical Notation for "they...so slow"..............348

List of Plates

Plate Page

6.1 The Mingograph Record....................................213

6.2 Mingograph and Movement Traces: Example VI.1a......219

6.3 Mingograph and Movement Traces: Example VI.1b......224

6.4 Mingograph and Movement Traces: Example VI.1c......231

6.5 Mingograph and Movement Traces: Example VI.2.......253

6.6 Mingograph and Movement Traces: Example VI.2-1.....257

6.7 Mingograph and Movement Traces: Example VI.2-2.....260

6.8 Mingograph and Movement Traces: Example VI.2a......263

6.9 Mingograph and Movement Traces: Example VI.2b......267

6.10 Mingograph and Movement Traces: Example VI.2c......274

6.11 Mingograph and Movement Traces: Example VI.3a......305

6.12 Mingograph and Movement Traces: Example VI.3b......309

6.13 Mingograph and Movement Traces: Example VI.3c......315

6.14 Mingograph and Movement Traces: Example VI.3d......324

6.15 Mingograph and Movement Traces: Example VI.3e......333

6.16 Mingograph and Movement Traces: Example VI.3f......334

Chapter 1

Introduction

1.1 Background To The Study

Until very recently the study of "talk," the study of language in general, has largely ignored context and contextual constraints. Researchers have generally disregarded the fact that the extraction of language from its communicative surround, for detailed analysis in isolation, is an artifice. However, if our intention is to offer description and analysis that move beyond the level of the trivial, it has become evident that the study of language must turn to consideration of language as part of the communicative context from which analysis has taken it. New questions are being posed and new problems are being raised that seem to be beyond the scope of traditional models. The data alone demand that research move beyond the arbitrary syntactic unit upon which modern language and language-related study has been founded and upon which it founders, that is, the sentence, to something else which has to do with semantic as well as syntactic considerations, and which takes into account the biological foundations of language. A truly scientific model of the process of human interaction must deal with the entirety of the communicative process, placing language in the context of natural spontaneous interaction, the ephemeral "ever-advancing now" which it structures and by which it is structured.

The early and perhaps necessary focus on language, the focus on words, arbitrarily and artificially extracted from communicative context, has been an extremely useful and productive one. Modern linguists have developed a powerful theoretical monument, with scientific rigor that continues to be the envy of other anthroponomic sciences. Modern linguistics has developed theory to provide powerful accounts of phonological and syntactic structures in language. Chomsky and his fellow workers' formulations of the Transformational Grammar (TG) model have apparently done for syntax what Sapir's "pattern principle" and the concept of the phoneme had done for phonology. Both provide descriptive and explanatory devices which organize discrepancies in previous models. Yet, as Chomsky and his colleagues developed their model, attempting to substantiate TG claims about the structure of language, it became clear that in order to account for Chomsky's syntactic structures we had to appeal to what to that point had been considered extra-linguistic factors. The Chomskian model and recent adaptations address syntactic relations, but Chomksy's theoretical statements, like the 1965 invocation of meaning, the claims about mental structures, and the appeal to localisation of brain function in connection with grammatical function, moved beyond syntax, posing some new questions that could not be handled by any syntactic model available. Bauman and Sherzer (1974), for example, speak of "something" that has been missing from the study of language, something that conventional linguistic research has not and perhaps cannot examine: patterns and functions of speech, the communicative context of discourse. Gunter (1974:21) also speaks of a "difficulty" inherent in much of the study of language:

> The difficulty, the thing that blinds us in our conception of such matters, is our rigid focus on the sentence and its parts. Perhaps this focus has been necessarily prior, but once having understood the formation of English sentences, we can turn our attention to context grammar, which embraces the features of English that bind sentences together into coherent speeches and dialogues, the real uses of languages.

And so for more than a decade now, linguistics as a discipline and those whose interests are bound up in the study of language have been suffering from a growing sense of anomie. Some feel they may have come up against the limits of the founding paradigm of the discipline.[1] In Kuhn's (1962) terms, we have passed the stage of "normal" linguistic science, the era of the 1950's and 1960's, and have entered a stage of anomaly. This stage is typically characterized by fragmentation and searching within and beyond the paradigm for new, more satisfactory perspectives. Linguistics has been no exception. Insoluble problems plague the central theory. Questions are no longer posed solely from within the paradigm, to provide missing bits of an incomplete model. Many linguists have attempted to "patch up" the model with reference to contextual constraints. Others have endeavored to incorporate insight from diverse disciplines into the paradigm. Some have proposed that semantic structure may underlie the production and use of language, and have sought to fit this notion to the TG model. Semantics has yet to provide an acceptable answer, however. In fact, semantics has engendered a split within linguistics itself. Interpretive and generative semanticists offer opposing interpretations of the interrelationship of semantics and syntax. Maclay (1971) briefly outlines approaches to semantics within linguistics. "Interpretive" semanticists insist that semantics must have input of syntactic information, and syntax must have input of semantic. "Generative" semanticists, on the other hand, propose that formal syntax must be analytically independent of a prior semantics.

---

[1] Perhaps not even most researchers from formal linguistics would agree to a "growing sense of anomie" of these proportions. They remain satisfied with what they are doing, content in their efforts to patch up the standard Chomskian model. Impetus for change, the growing sense of alienation and frustration seems to come not so much from the core of linguistics proper, but from the "lunatic fringe" of linguistics and from those outside who have an interest in the study of language and language-related phenomena. It is primarily these latter groups who have expressed dissatisfaction with and feel constrained by the standard theory. And they have had their effect. In spite of the tenacity of the traditional core, it is possible to document a move throughout linguistics away from context-free Chomskian grammars toward a study of language as part of a larger communicative context. The concern is now with a "science of mind."

At present, then, there is no coherent paradigm for the study of communicative interaction. A number of disciplines address the issue, among them linguistics, anthropology, sociology and psychology. Each has its own tradition, its proclivity for a particular aspect or paradigm in the study of language. All too often each school, even each book within a school, sets out a new and different paradigm for study, defining not only the subject and the questions, but the attempts to answer them. A plethora of various, differing, incomplete and confusing theories and postulates must be mastered to access current debate. The study of human interaction is a jungle of conflicting and contrary notions and theories. Each school, each postulate, demands acceptance of and allegiance to a particular definition of what is to be considered acceptable and appropriate in terms of data, questions, answers and so forth. What is more, the several theoretical frames, as well as the units and methodologies they employ, seem mutually exclusive. Such mutual exclusivity makes for hard choices, the necessity of which seems at least questionable. As Mair (1977:X-3) argues, "[if] the categories will not mutually translate, then one must choose and then one's research will be partisan, one's questions posed by a decision to 'join' a way of thought." Each conceptual frame in the study of human interaction and/or language thus artificially delimits and confines the scope of the enquiry. As a result, at least one researcher has chosen to disregard the competing claims and to base his argument about "mental process" (in the process of human interaction) on observable data only.

1.2 Mair's Model of the Melody of the Text

The work of Michael W. Mair is an attempt to "get at" the reality of the process of communicative interaction. His is an eclectic approach, one which brings in elements from many diverse and competing areas of study. Rather than pledge allegiance to a particular school, discipline or paradigm, Mair has presumed to trespass upon all. His hope is that an eclectic approach might, in refusing to deny the relevance of data, theory, argument, on paradigmatic or disciplinary grounds, begin to get a handle on interactive reality. A broad scope may facilitate focus on that which is studied, rather than the school which performs the study.

In his own words, Mair's decision has been "for the study":

> After it was clearly realized that the state of play within and between disciplines studying human interaction was a jungle, a decision was made to disregard Medawar's admirable dictum [i.e., Science is the art of the soluble; clear answers can be achieved only for clearly defined questions] for a while, and with these conflicting theoretical awarenesses on board, just to look at the physical record (videotape and laryngograph signal) again, and again, and again, in the hope that new and perhaps important questions and answers might show up simultaneously. The experimenter hoped that his own brain might, with sufficient exposure, act as a kind of filter for the systemic principles in operation during ordinary conversation, that they might slowly become evident (1978:20).

The outcome of this lengthy process of observation and reflection is the set of "mutually dependent notions" of regulation of interaction put forward in his Steps Toward Principles of Text Regulation (1977, 1978). They compose a unique construct in the study of human interaction. It is this set of "mutually dependent notions" to which I appeal in the present investigation.

1.2.1 Approaching Mair's Model

Mair is an opthalmologist with training in anthropology and neurology. He comes to the study of interaction from what are flippantly called the hard sciences. Mair expresses initial dismay at the chaotic, "pre-paradigmatic" state of his chosen field of study. By contrast, the hard sciences maintain a generally accepted paradigm which informs theory, methodology, and definition of acceptable data. At bottom there is what Mair calls the natural science model: objectivity and the observable are ascendant. Questions come from the paradigm. Questions deal with problematic areas of the central theory; they represent efforts to provide missing "bits" of an incomplete model.

Mair presumes upon what he calls natural science and natural science variables to steer him clear of the methodological, theoretical, and epistemological quagmire which surrounds the study of human interaction. He aims to avoid the pitfalls that have plagued the largely impressionistic studies of human interaction in the past. "If the data is not (sic) Natural Science variables, then it comes from some theory, either explicit or implicit, about behaviour. Common sense is just one such theory," (1977:IV-2). Mair's goal, ultimately, is to establish a true Natural Science of human interaction. "The founding assumption of the project is that to establish a true Natural Science of human interaction is possible and desirable," (1977:IV-1). He is anxious to establish Natural Science, i.e., objective parameters for the description and analysis of interaction. Mair intends to relate and account for a "natural segmentation" of language, to "consider the way that the sensory modalities were patterned by motor activity during interaction" (1977:IV-11). His particular concern is to account for spontaneous interaction in the dyad.

Mair's research is thus directed by certain a priori dictates, the epistemology of the natural science model. At the same time, Mair is very concerned with the problem of "over-objective" studies, as described in his own work (especially 1977, 1978) and in the work of Scollon (1981a, 1981b) among others. The problem has to do with identification, definition, and measurement of units and the definition of boundaries between units. The problem has been framed many ways. Mair himself presents several versions. Perhaps a most accurate and economical portrayal takes form in his discussion of the "prescriptive" versus "descriptive" dichotomy. That is, are the units really "out there" (descriptive) or are they an artifact of the perceptual, verbal, or cultural categories of the observer (prescriptive)? Are we talking about a physical sequence of events or about descriptive labels? The notion of "fact" itself is at issue. That is, anatomical description of body movements and/or acoustic analysis of the speech stream tells us nothing of the significance of the movements or aspects of the acoustic stream which we have isolated. However "objective" they may be, artificially extracted units may have very little to do with what is studied, what is really going on.

Mair remarks that there are several "entities" in the literature "meant in some sense or other to be manifest in interaction" (1977:X-2). There are, he notes (1978:28-29), Bernstein's restricted versus elaborated codes; Chance's agonic versus hedonic modes; Smith and Apter's telic versus paratelic modes; Argyle's notions of intimacy versus distance; Ekman's concept of emotions, emblems, illustrators and adaptors; and Birdwhistell's hierarchies of kinemes and kinomorphs. These "entities" serve not only to document the fragmented nature of the effort in the study of interaction but point up a deeper dichotomy underlying communicative interaction research. This deeper dichotomy has been variously called "expression versus communication," "verbal versus nonverbal," "emotion versus intellect," (Mair 1977:X-2). Mair submits that to choose either aspect of the dichotomy is to pledge allegiance to a particular epistemological frame and is to prejudice unfavorably the theoretical frame of one's research. He refers to a comparable problem in methodology: "Related to the debate about dichotomies is the long-standing methodological problem about description versus explanation, imposed versus derived categories, natural units versus units of convenience. Ultimately the opposition is between words for experience and the experience of words" (1978:21).

From this point, Mair argues that the attempt to put labels on social, physical, or physiological processes always results in a cultural interpretation. He refers to the studies of Ekman, Argyle, and Birdwhistell which indicate that all our notions of, for example, "body," are culturally mediated. Words for feelings and emotions demarcate sectors or experience which owe their identity to the "synchronic cultural scheme" of which they are a part. Mair proposes that appeal to nature for validation, an appeal to the common sense "that's the way things are," which he argues is one of the few cultural universals, depends upon the cultural categories available and may in no way pretend to account for why the categorization of the world takes that particular form. As Eibl-Eibesfeldt (1967) notes, even the most rudimentary description embodies an inherent bias, an interpretation (implicit or explicit) of the nature of "what is out there." Complete description demands, is, nothing less than complete replication. But description itself requires selective chunking of the continuum which in fact is the reality. We are caught:

> I believe that one of the brain's own properties gets in the way of felicitous description. The brain makes categories. . .[isolating] discrete entities in what otherwise would have been a continuum. . .[However, if] I do not create such isolates, then I cannot talk to you at all! Such are brains (Mair 1980:6).

Mair submits that we can formulate a natural science of how it happens, but "questions of 'why' or 'what'--for example, what is an emotion?--are not answerable except in culturally satisfying and self-fulfilling language (1977:IV-10). An express goal of his enterprise is to establish a "proper relationship" between the social and natural science "understandings" (1977:X-2). He aims to find a heuristic connection between cultural concepts and natural science facts.

> The relationship between descriptive labels for both stretches of speech, particularly the intonation or tone of voice, and faces in action on the one hand, and the physical sequences of events in both voice and movement on the other, is a central pre-occupation of the theory (1977:X-4).

1.2.2 Early Research Into Communicative Interaction

In this endeavor all workers have had and continue to have the same fundamental problem: validation of chosen concepts and units (see Mair, 1978). Mair insists that we must clarify the exact nature of our investigation, defining appropriate, acceptable questions and methodologies.

> The first task was to search for questions, not answers because [following Medawar's dictum] it was felt that particular answers can only be in response to particular questions. So first it was necessary to study the idea of questioning (1977:V-5).

Mair found that much of the early research on human interaction depended upon interjudge reliability and interobserver agreement for identification, definition, and description of units in speech. Interobserver agreement served to validate chosen units which themselves seem to have originated in some common sense notion of the way things are, coming, in Mair's terms, "out of the top of the theorist's head, in fact." Pause, stress, and timing are three of the units selected in this manner by early researchers. Nonverbal communication studies tend to evidence similar assumptions. Identification of units is not generally regarded as problematic. Researchers have been predisposed to focus on "obvious," "easily identifiable," and stereotypic body movements. Behaviours are by definition held to particulate, i.e., relatively simple, components that appear at the end of a variable sequence of events. Generally, units of behaviour have been put forward as discrete, repetitious patterns with specific function(s).[2]

Such a complex of notions incorporates major theoretical and methodological flaws into the study of communicative interaction. Mair's remarks are to the point. He maintains that the discrete nature of units of behaviour and units of acoustic analysis may be, above all, an artifact of the process of observation by observers:

> I would argue that the objective way these behavioral units are identified merges into the observer's decision-making as one ascends the evolutionary scale; and that inter-observer agreement, however carefully calculated statistically, might only be a measure of the precision of culturally coordinated decision making. (1978:7).

Thus even trained observers are biased by their epistemological a priori, by the training which seeks to make of them objective, expert, "super-observers." As Osgood (1966) admits, studies relying on interobserver agreement may reveal more about the semantic systems of the observers than anything in the data. Furthermore, the units chosen as significant by the observer may not be those that are significant to participants. But, it would seem generally, and especially in the study of interaction, that the only truly relevant units are those which have significance and value for participants; i.e., those units which effect changes in and among interactants themselves. Mair argues that the need is for an empirical, scientific, methodology to determine what the real units are. "It will not do to rely on intersubjective judgements. . .It is precisely how such judgements are made that we are trying to investigate," (1978:9).

---

[2] The notion of behaviours as discrete entities seems to derive from the notion of "fixed action patterns" attributed to birds by ethologists.

1.2.3 Recent Interaction Research

In recent years researchers have demonstrated an increasing concern for the objective validity of their work. They have consequently made large steps toward objectification in the study of communicative interaction. The use of precise mechanical devices and computer analysis has become commonplace. However, objectification and mechanical precision have not eliminated the problem of definition and identification of units. It is not at all clear that what is being so finely measured, what is being studied, is of any major relevance to the process of communicative interaction. Once again, Mair's (1977) remarks are to the point. Mair submits that the limits to accurate, objective measurement are the limits of technology, but what we must search out are the perceptually significant cues for interactants, the cues upon which participants depend for interaction to occur. We are missing something yet.

1.2.4 Mair's Model

Mair attempts to get at that something. Units in Mair's model are defined by interactants themselves. Identification, definition, the constitution of units comes from exhaustive study of spontaneous dyadic conversation. Chosen units are those which are found to be perceptually significant for interactants. Such units are, Mair argues, evident from study of the observed changes in behaviour (taken here in its broadest sense, to include both verbal and non-verbal production) produced in one participant by the other. From this perspective, the initial task is to identify those changes which are crucial to transform one percept into another. Then perhaps we can begin to reconstruct the mechanism, the melange of sound and movement which apparently guides topics to their conclusions in conversation.

To begin, Mair assumes that there is something in the acoustic stream and physical movements generated by interactants which guides topics to their conclusions. Following on from the long debate in linguistics on whether categories of linguistic theory are "descriptive" or "explanatory," "God's truth" or "hocus-pocus," Mair proposes that there are in fact patterns "out there," in the visual and acoustic modes, and that interactants use them to generate the process we know as interaction: "[I]n sound and movement--air patterns and visual pattern in time--there must be an order or structure which is really there, i.e., God's truth, and a way to discover it might be to see how sound and movement cross-cut in interaction" (1977:IV-11).

In this way Mair aims to achieve his natural segmentation of language, reconstructing the way in which sensory modalities are patterned by motor activity during interaction. Mair argues that if we are to achieve a natural science of interaction, we must look at the patterns developed and employed by participants as they interact. We must begin to think in terms of systems. "The order--Natural Science order--of communication might. . .be best studied by examining what happens 'when two or three are gathered together,' not one pontificating in a void," (1977:IV-8). Mair submits that the need to examine systems, patterns, rules is particularly pressing to the study of communicative systems: communicative systems, unlike many other naturally occurring systems, are not reducible to natural science variables. "The significant point about communication systems, as distinct from the others, is that they 'take off' from the natural science base, and that the explanation of items within them are irreducible to natural science units" (1977:II-8).

The notion of system as emergent from the organism that manipulates it is one of two key notions which Mair borrows from formal linguistics. (The second has to do with the "meaning," or lack thereof, attached to individual units of the system.) Early in his career the prominent linguist Noam Chomsky documented the imperative to look at language as a system of rules and patterns. He saw language as formal, an abstract system of "elements" related by grammatical rules of production and perception, i.e., syntax. The system of language was conceived of as an entity separate from either its users or uses. Chomsky provided supporting evidence for his notion of language as system in his study of child language acquisition. He was able to demonstrate that the key in child language acquisition is rule acquisition, and not, as previously thought, accumulation of successive reinforcements for each language item.[3]

The notion of language as "system" is closely related to the second borrowed notion: individual or minimal units of language are "premeaning." That is, formal linguistics maintains that units must occur in relation to one another, specifically in grammatically acceptable configurations, for them to take on meaning. On their own, isolated, units are meaning-less. It is the system, the rule and pattern governed configurations which resolve the units into meaning.

---

[3] In the process, Chomsky exposed the fallacy of the fundamental assumption of behaviourism--that truth could be accumulated through examination of single variables.

> In the end, it is evident that there must be 'something in' the acoustic or kinaesthetic stream which interactants, animal and human, can find structure in, according to shared rules, such that common action occurs. That these units and rules are hierarchical seems established, and this implies that the individual units are at least 'pre-meaning,' if not actually 'meaningless,' (1977:I-9,10).

Mair proposes to call the "unit" of sense in his study of human interaction the "model." He defines "model" as "a realisation of a state of affairs" (1977:V-7). The proposed model is both a physically demarcatable section of the speech-with-movement stream and an information unit. Mair makes mention of several comparable notions in the interaction literature: O'Connor (1974) investigates the "word group;" Goldman-Eisler (1968), among others, works from the "sentence;" Kendon (1967) proposes the "tone group;" and Thom (1975) posits the "act." "Model," then, is a broad term meant to encompass more than merely vocal production.

Mair submits that "what is communicated" may be most economically and accurately described as a "model" or "state of affairs." He maintains that a primary problem in definition of suitable units for the study of interaction comes from the fact that for communication to occur, nothing at all need actually be said. A shrug, a wink, a gesture, a chuckle, an "um" or an "ah," even silence--which, after Birdwhistell (1970) can be "so loud as to drown out the scuffle of feet"--may be sufficient. Thus the notion of "model." Mair identifies two features of models: models are implicit (to a greater or lesser extent); and models are "shared," i.e., people are immersed in them rather than exchanging them like counters. "We call these models 'interactional' to stress that in colloquial conversation, the model is a joint production, and evolves by the collaboration of the participants" (1977:V-7).

Mair describes his investigation as a study of "social physiology." The term is meant to account for something observed in interaction, that which he terms a "direct wiring mediated by the senses among individuals" (1977:IX-3). In interaction, he says, both (or all) parties are "together in time." Mediated by the senses, the brain of one acts directly upon the other. This is how the participants are directly "wired into" one another. The notion comes from the sense of touch. In touch, Mair points out "the actual tactile patterning of the sense organs at the point of contact necessarily has the same time form for both parties" (1977:IX-3). Mair is particularly concerned therefore with what has often been called "neuroplasticity," the inability to separate neurological function from its environment. The observed ability of the brains of individuals to take the timing of their own cognitive processes from rhythms impinging on the senses is of particular interest to the study of human interaction.

Mair posits a "social brain":

> The idea is that even before social patterning has occurred; the individual psychology is incomplete,

> only part of a system of which other individuals possess the other parts. The fact that this social brain is not joined up by actual nerves but by the sensory modalities has obscured for us that the system is supra-individual, that to understand it we must study not what one individual possesses, but the patterns set up among individuals when they interact (1978:12).

With these notions Mair obviates at least one long-standing "need" in formal linguistics, one which has often prejudiced research in the past. Briefly, Mair's claims about social physiology and a social brain abrogate the need to find specific centres in the brain which correspond to entities demarcated by categories chosen by researchers to describe and analyze communicative interaction. Chomsky's claims about mental structures and his assertion of the connection between grammatical function and localized brain function are but one manifestation of this "need," what Mair maintains is a wrong way to approach the study of interaction. We must, he argues, look at patterns of interaction; we must begin to think in systemic terms about "rules." He proposes that the imperative is to examine the dynamics of human interaction systems as they are employed, in natural spontaneous conversation.

1.3 The Present Study

The present study replicates some important aspects of Mair's work but it is not a replication in a formal sense. The general object of the present study is to try Mair's explanatory framework with respect to some segments of collected speech and interaction data, in order to demonstrate its explanatory adequacy and to suggest some of its limitations. One aspect of his work that is consonant with the work of several other investigators is that of synchrony, what has been often described as either complementarity or simultaneity of sign. The differing notions of that term (synchrony) are developed in subsequent chapters of the thesis, and the choice of the term, synchrony, as a common, unifying theme in several authors' works, should be made clear after a review of literature.

In this study I address three half-hour conversations as representative of interaction generally. That address is further refined to circumscribe no more than one brief fragment from each interaction, including perhaps thirty seconds of real time data in all. The three fragments selected are taken as represenative of a certain common and continuously recurring genre of interaction, one under which conditions of "rapport" obtain, and in which a particular kind of sound and movement behaviour--synchrony--may be found to occur. There are similar fragments throughout each half hour interaction. The specification of three ten-second fragments as interactions of interest allows for a kind of detailed analysis that an appeal to a range of interactions as minutely analyzable data source would make cumbersome and, in the end, unmanageable.

It is not a unique claim; there is ample precedent for it in the literature. The work of Labov and Fanshell (1977), Urion (1978), Condon and Ogston (1967), to name but a few, consistently specifies a narrow data base and advances general claims about a class of events based on analysis of that specific data. As Rene Thom has pointed out, there are at least two ways to study a phenomenon: to collect many examples and generalise--that is, to study the "species" characteristics; or, to take just one example and analyse, and then see if what one has discovered holds true for other examples. The present work follows the latter course, invoking a long tradition of microanalysis in the study of human communicative interaction. Finally, I appeal specifically to the procedures established by Mair in his original investigation of the process of human interaction.

The data for this study come from audio and video recordings of conversations, each approximately one half hour in length, that took place over a two-month period from November to December, 1980. The field work involved, if in fact there can be said to be field work in a study such as this, transpired over the preceding year and a half. During that time I was privy to the research project of a graduate student in the Department of Educational Psychology, University of Alberta, which was similar in many respects to the present study,[*]

---

[*] The project in question involved videotaping in split screen sessions during which a psychologist administered WISC-R tests to individual students. Videotaping took place in the Group Processes Laboratory in the Department of Educational Psychology, the room which was later to be used to collect data for the present study. Complete transcriptions were made of verbal interchange and intonation, as well as notations of head and body movement. Description focused on large body movement, eye contact, and intonation. Description was couched in impressionistic

"walked through" the audio and video record of the half hour conversation between a male and a female, the freshmen at Cambridge, that forms the primary data base from which Mair developed much of his model; performed three preliminary and experimental recordings of participants in dyadic interaction; and was party to numerous casual sessions where the parameters of Mair's model were informally evaluated *in situ* (e.g., in restaurants, shopping malls, social gatherings) as to observational, descriptive and analytic adequacy. My attendance at all these events was not as participant or interactant, but as researcher, observer, and data collector.

Within the framework of anthropology outlined by Hymes (1977) these enterprises might be said to constitute a range of field experience and field work. Attendance, participation, and observation in the various interactions of everyday life constitute a kind of field experience, as does the exposure to the two bodies of data. Those events wherein I attempted systematic categorisation of descriptive parameters, either formally with recorded data or formally *in situ*, constitute field work. Microanalyses of audio and video recordings contributed to field work, and were in fact additional trials of observational adequacy. In anthropological terms the present study might also be

---

[4] (cont'd)terms. A hand-drawn representation of perceived intonation contours formed an essential part of the project. S-shapes were of particular note. The project was designed to investigate some of Mair's claims about the process of human interaction.

described as an ethnography of sound and movement behaviours in communicative interaction. Such description insists upon the interrelationship of the two modes of research activity defined as field work and field experience.

The nature and purpose of the present study dictate within specific and explicit limits what are to be considered acceptable data, and, consequently, severely constrain means by which data are collected. Microanalysis of sound and movement through short segments of interaction requires a precise and finely detailed data base. To date, such data are obtainable only within the confines of a laboratory setting. "Candid camera" techniques are not applicable; they cannot as yet provide adequate, sufficiently accurate data. The laboratory setting obviates many of the problems that traditionally plague anthropological and educational research. The roles of participant and observer, which tend to merge and blend in common research situations are more clearly defined, made explicit by the fact of audio and video recording technologies and techniques. The laboratory setting demands that all parties be aware of the primary purpose of the meeting, i.e., data collection. There is only one possible reason for the presence of the researcher, the microphones and cameras, or indeed for the interaction itself in this setting.

The researcher is "participant" in that he must instigate the interaction; he must choose participants, arrange time, place, setting, and acquaint subjects with their surroundings and each other in such a way as to facilitate interaction, and especially in this case, the creation of "rapport." But when the microphones are on and the cameras rolling, it is the subjects themselves who must create, in time and from the resources at hand, what will appear on the audio and video recordings as boredom, tears, bewilderment, delight, understanding, disgust, in sum, human interaction. The other observes. During the present study, as researcher, I was always present, available and accessible throughout the interaction. On occasion subjects spoke of me; once or twice they spoke directly to me, usually in regard to a technical matter, but for the most part as the interaction progressed I was dissolved into the background of tables, chairs, microphones, and cameras. I was a technician, attending to the machines, and an observer, collecting data. It was their show.

A major methodological problem involved boundary definition. That is, how was I to specify beginning and end points for the fragments of "text" (conversation) selected to represent the best, most complete evidence of synchrony as an identifiable, observable, definable aspect of human interaction? There were no apparent unit markers. A seemingly obvious boundary marker, the "pause," proved problematic (as expected). "Pauses" that appeared to be evident in the progression of aural and/or visual modes, i.e, moments of silence or hesitation, often proved under close inspection to have amorphous, ill-defined boundaries that extended over two or three tenths of a second in time, well beyond the acceptable time scale of this undertaking. What is more, the problems encountered in attempts to identify and define "pause" in the speech stream, as discussed in Chapter 2 in relation to the work of Lehiste, Huggins, Ohala, Scollon, et al., are compounded, magnified many times over, when consideration is expanded to include events in the visual mode.

For example, in looking at human interaction one must account for the fact that one mode often overlaps with, and carries on from, the other, continuing or completing the state of play being delivered with their occurrence. Movement may, and does, continue across silence; speech may be sustained through hesitation. It was not at all clear from the data, then, that "silence" or "hesitation" could in fact be construed as boundary markers at all. Placement, occurrence seemed contextual and did not always indicate a beginning or ending. Communication carried on across silence, at times using silence to help deliver the state of play; and movement, gesture, appeared to continue through hesitation. What is more, in both visual and aural modes the previous is shaped by preparation for the next, and the succeeding carries on from the preceding. Meaning does not occur in discrete units, but rather movement and voice blend and merge, shaping one another through time. Thus the apparently obvious disappears into a continuum of sound and motion.

It was thought that semantics and/or syntax might offer a way out; that grammatical units or formally-defined units of meaning might be so circumscribed as to allow specification of criteria for indication of boundaries, and subsequent definition of boundaries themselves. That hope, too, was generally frustrated. Utterance was regularly ungrammatical, and grammatical units, where specified or specifiable, often disagreed with semantic specification of "unit." Focus on "meaning" engendered a myriad of difficulties that began with "what is meaning?". Furthermore, although often clearly demarcated and sharply defined in time, broader semantic categories such as "topic" proved unsuitable as boundary markers for the present study as well. Topics tended to occur over minutes, not seconds; and smaller units of meaning were particularly difficult to identify: they were inclined to overlap, merging into one another.

In the end, the limits of technology proved to be the influential factor in what was to be a rather arbitrary decision. First, the transcription equipment (a mingograph) delivered a decipherable trace of the fundamental frequency of voice only when paper speed was set a 100mm/second. The length of mingograph traces from fragments in excess of 10 seconds real time was prohibitive. Mere display of all the data on that scale was virtually impossible. Second, through each fragment of text there are, for each tenth of a second interval, three data points for description of physical movement. That means that in, say, a 10-second fragment there are 300 data points for movement alone which must be transcribed, described, evaluated and analyzed. In addition, the purpose of the study demands that movement and sound be evaluated in terms of each other, through time. Thus, the sheer magnitude of the data sample from extended fragments of text must produce a data base so bulky as to be unmanageable. These were strong arguments for minimizing fragment size.

A third factor, problematic in itself, proved decisive in definition of selected fragments. In order to make claims about synchronies of sound and movement it is necessary to be able to compare manifestations in both visual and aural modes accurately to within 0.1 second. It was therefore imperative to match the mingograph trace to the visual transcription. Calibration problems, primarily in the form of disagreement between time codes in the audio and video records, complicated an already complex task. A decision was taken to use "silence," pauses in the speech stream, as markers with which to match audio and video records. "Silence" was identified first subjectively by ear from the video tape, and corroborated on the mingograph trace. Silence markers in turn were arbitrarily taken as boundary markers, defining and delimiting fragments selected for microanalysis. In this manner, selected fragments of appropriate length, roughly 10 seconds in duration, were clearly defined and located in time on video and audio tapes, and in mingograph traces.

It is perhaps important to note here that trial analysis of segments in excess of 10 seconds demonstrated that more lengthy sections of text delivered only "more of the same." The evidence for synchrony in the selected fragments of shorter duration was strengthened primarily by repetition in longer fragments. The essential evidence seemed to be all there in the shorter sections selected.

Finally, it is imperative to specify explicitly the criteria by which I categorize in this study. As participant in innumerable conversations like the ones recorded for analysis, and as participant, albeit at one remove, in the interactions of interest, I make the claim to know how to comport myself. That is, as a social human being I make the claim to be able to interact successfully, to "know how to do it." As researcher, I observed the events of the interactions of interest with an end in mind other than, in addition to, that of making distinctions along any number of idiosyncratic bases which would allow and result in what is commonly called "making sense" of the events of the interaction in personal terms. I was constantly evaluating what I perceived to be the "reportable" flow of events in the terms proposed in Mair's model of the principles of text generation. I made video recordings of the visual events, and from them, two-dimensional transcriptions of face and hand movements in 0.1-second intervals of selected fragments of each recorded interaction. I made audio recordings of the acoustic events of each interaction, and from them, extracted graphic plots of fundamental frequency for selected fragments, marking 0.1-second intervals in time. I matched audio and video records to within 0.1 second, and compared manifestations of sound and movement within and between interactants in time. I made additional notes on features deemed notable. I entered taping sessions and began examination of recorded data with a priori definitions of categories and parameters for specification of categories in the process of human communicative interaction.

The basic question for the study was how to document the existence or non-existence of synchrony in its several and various forms as an observable, identifiable "event" in human interaction. In other words, using the criteria from which Mair had originally suggested his final "artificially-extracted parameter" to engender description and analysis I attempted first to identify and then to explain and to account for the perceived and intuitively obvious instances of intra- and inter-personal synchrony manifest in the audio and video record of three interactions.

1.4 Organisation of the Thesis

Chapter 2 reviews some of the pertinent studies in discourse and prosody.

Chapter 3 is a review of Mair's work and Chapter 4 is a critique of it.

Chapter 5 details the specific methodology of the study.

Chapter 6 includes the presentation and interpretation of data from the three fragments of selected speech and movement interaction.

In the final chapter, the implications of the study are summarized and suggestions are made for further research.

Chapter 2

Current Studies in Discourse and Prosody

2.1 Speech Act Theory

Although there is no coherent paradigm for the study of communicative interaction there are some common threads running through a good deal of the relevant research. The concept of the speech act as it has been developed in philosophy by Austin (1962) and Searle (1969) has, for example, been integral to a number of recent efforts to describe communicative behaviour. The concept of the speech act grew out of Austin's observation that "not all sentences are indicative, i.e. describing a state of affairs which is either true or false" (in Vanek 1979:274). Speech act "meaningfulness" refers to the meaning found both within the system of syntax and within a system of acts. The speech act is held to have three properties: locutionary, or the act of speech as speech; illocutionary, what is done in the process of speaking, or what the act represents; and perlocutionary, what is effected by speaking. Urion (1978:26) notes that these latter speech act properties require description in "other than linguistic" terms.

Fillmore (cited in Urion 1978) developed an analogy from Searle's theory of the speech act:

> Syntax. . .characterizes the grammatical forms that occur in a language, while semantics pairs these forms with their potential communicative functions. Pragmatics is concerned with the three-termed relation which unites (i) linguistic form and (ii) the communicative functions which these forms are

> capable of serving, with (iii) the contexts or settings in which those linguistic forms have communicative functions. Diagramatically,
>
> Syntax [form]
> Semantics [form, function]
> Pragmatics [form, function, setting]

Linguists have in this way come to use the term "pragmatics" to refer to the social correlates of speech use, and, to describe the social-contextual constraints of language use.

Many linguists, Fillmore, Ross, Goddard and Saddock among them, have developed frameworks which include these conceptual constraints and/or considerations in formal grammars of discourse. Sociologists working in the area of conversational analysis--Goffman, Sacks and Schegloff, for example--employ as a basic unit of description the concept of utterance as qualified by Searle in speech act theory. Hymes' work in the ethnography of communication makes reference to the speech act, not as a component in his descriptive framework, but as the minimal unit of description. More recently, Labov and Fanshell (1977) make use of speech act theory in their explication of dyadic interaction.

The problem with these approaches, the problem with speech act theory and its offshoots, lies in the unit of description. The locus of analysis is still the sentence. The sociological concept of utterance is an equivalent. Labov and Fanshell, for example, claim "despite their citing and dealing with utterances that seem, temporally, to predict other utterances, that there is still no necessary coherence between utterances" (Urion 1980:511). The level of explanation has not extended beyond the sentential. Such research is unable to move beyond the arbitrary syntactic unit upon which it is based.

2.2 Acoustic Phonetics

Another tradition of import is that coming out of acoustic phonetics which focuses on quantifiable analysis of temporal structure in speech. It is evidenced in the works of Pike (1945), Kozehvnikov and Chistovich (1965), Huggins (1975), and Ohala (1975), among others. This tradition has also been marked by the constraints of the arbitrary syntactic unit upon which it is founded.

Ohala (1975) offers an excellent description and analysis of the state of the art. He documents three simple, basic hypotheses in the literature:

1. Some units of speech, perhaps syllables or "stresses," are uttered in time to an underlying rhythm. This is the Isochrony Hypothesis (IH) (Huggins 1975).
2. The units of speech are executed according to some underlying preprogrammed time schedule although there may be no Isochrony in this schedule. Ohala (1975) christens this the Comb Theory.
3. There is no underlying time program or rhythm; any given speech event is merely produced after the preceding event(s) have been completed. One "unit" is strung after the other. This is termed the Chain Model (Ohala 1975).

2.2.1 Isochrony Hypothesis

Proponents of the IH propose a metrical foot as a unit in either the production or perception of speech, or both. A foot is formally defined as a sequence of syllables containing one and only one stress syllable. There is usually a further proviso that a "foot" must begin with a stress syllable. This definition of foot nicely specifies boundaries that might otherwise require lexical or syntactic specification, but has led to the creation of an imaginary "silent stress" syllable to account for and legitimate any initial unstressed syllables. The IH suggests that "feet" are of equal duration, the product of underlying cognitive processes or rhythms. Lenneberg (1967) for example posits a 6 Hz underlying rhythm for human speech.

Uldall (1971) also found evidence of isochrony. She analyzed an audio recording of a 45-second reading of "The North Wind and the Sun," read in a "moderately slow, news reading style," by Professor D. Abercrombie. Dividing the text into rhythmic segments on the basis of syllables marked "strong" by the speaker-listener, Uldall found in the text a tendency towards the the production of isochronous metric feet. In a similar study, Ohala (1975) examined one and one half hours of oral reading from technical prose. He proposed that sampling running speech and measuring the intervals between successive release of an "easily detected" speech event associated with syllable onset might test the existence of isochrony in speech. "If there is some kind of periodicity underlying speech these intervals [between speech events] should coincide with the basic period of this [isochronous] rhythm" (1975:432). Using a computer, Ohala marked peak jaw openings as his easily detected speech event associated with syllable onset. He found no multimodal distribution, no multiple peaks of equi-distance, no evidence for underlying rhythm; in short, he found no support for the claim that there is an isochronic principle underlying speech.

These studies are seriously and obviously flawed, replete with confounding variables that obscure the existence or non-existence of an underlying rhythm. Ohala's investigation used a marker that is not reliably correlated with any underlying neurological event; Uldall's analysis depends on a subjective demarcation of the text into "feet;" and, none of the studies investigated spontaneous speech, the nature speech of interaction, "as it happens."

In a further study, Ohala (1975) measured intervals between successive drops in oral pressure, "such as would occur upon the release of voiceless obstruants" (1975:434), in approximately one hour of spontaneous speech. He found possible evidence for the IH; there were apparent sub-peaks approximately 50 milliseconds apart in the histogram of his data, but they are "enveloped in noise and do not always seem to be spaced evenly" (1975:434). Huggins flatly states that the IH taken literally is wrong. However, on evidence presented by Barnwell (1971) and Lehiste (1973) which demonstrates possible isochronous tendencies in parts of the data, he concurs with Ohala's conclusion that, although there is no concrete evidence to support the notion of an isochronic rhythm in speech, it may well be "useful to keep looking for one."

2.2.2 The Comb and Chain Models

In a watershed study Kozehvnikov and Chistovich (1965) proposed that the validity of these two hypotheses could be tested through "a special statistical analysis of the duration of segments in a given utterance repeated many times by a subject." They argued that the two models should demonstrate differing relations between the variance measured in any given large interval and the sum of the variance of its component "sub-intervals." That is, in the chain model the variance of the large interval should equal the sum of the variance of the component intervals, while in the comb model the variance of the larger interval should be less than the sum of the variances of the component intervals. This effect should hold, they argue, because in the comb model variance error is shared by adjacent intervals. They are negatively covarying. Ohala summarizes: "in the comb model [it is proposed that] the error in any 'x' segment is shared by two adjacent intervals [and] will contribute to the error of both intervals in equal magnitude but opposite sign, thus making adjacent intervals negatively covarying" (1975:439). Kozehvnikov and Chistovich found evidence of negative covariance in their data and thus argued that their data provided evidence for the existence of some sort of time program "at least as long as the word and perhaps as long as the whole sentence." They found for the existence of the comb model in speech.

Lehiste (1971) found corroborative evidence for the comb model in studies of word lists of ten items repeated approximately 110 times by two subjects. The data demonstrated the existence of temporal compensation "between certain pairs of segments and within all the segments that constitute a word" (Lehiste 1971:168). Huggins' (1972) experiments document analogous temporal compensation, offering further proof that the comb model holds in the temporal structuring of language. And Allen (1969) found existence of negative correlations between adjacent intervals in the speech material he measured. But Ohala (1970) found for the chain model in his data, and Lindblom (1970), in a test of Chomsky's Nuclear Stress Rule could find no effect of one word on the duration of another.

Ohala (1975) proposes that although these findings seem contradictory they are in fact "compatible"--but they do not demonstrate whether the chain or the comb model applies to speech more properly. Kozehvnikov and Chistovich argued that the comb model could be accepted if the variance of the whole interval was found to be less than the sum of the variances of its component parts. They, and several other researchers, have found this relation to hold in their data.

Others have found an opposite relation: the variance of the whole is equal to or greater than the sum of the variance of the parts. Ohala submits that when a small sub-interval is chosen as the "unit" of analysis, we will find that the variance of the larger interval is less than the sum of the variance of its parts, as the comb model suggests; but when large sub-intervals are chosen, the variance of the whole will be greater than or equal to the sum of the variance of its parts, suggesting the chain model or some other alternative.

In addition, Ohala (1975) challenges the validity of the statistical procedures employed. He argues that the statistical analysis is itself invalid. Variation from repetition to repetition in the rate at which the subject utters the test sentence or words "contaminates" the data, as does measurement error, which may derive from "sloppiness," i.e., mis-measuring an interval, or from wrongly segmenting the speech data, i.e., not knowing how to segment our speech because "we do not know what the brain of the speaker considers to be an 'event'" (1975:441). And because we cannot control for or corret the latter source of error, statistical analysis is invalid. Such analysis confounds itself: measurement error tends to produce data which support the notion of the comb model, while rate variation may produce evidence for either the chain or the comb model. Thus confounding factors in speech analysis methodology may determine the results of the study, regardless of the validity of a particular model. Ohala's (1975:442) remarks are (again) to the point: "the confounding factors could completely determine the outcome one way or the other, regardless of which model holds. . .whichever [covariance] relation is obtained will not indicate whether the chain or comb model better accounts for the timing of speech gestures."

In efforts to disambiguate methodology and data researchers have resorted to other, additional, statistical methods: normalizing durations of the whole utterance, and expression of temporal variability of speech segments in terms of relative error. Ohala (1975) and Lehiste (1971) "normalized" by limiting analysis to utterances which had durations closest to the mean; Allan (1969) performed a multiplicative operation where large and component interval durations are subject to a normalizing factor which artificially equalizes their total duration. Nothing useful is gained by these techniques, Ohala (1975) argues; they merely alter the data artificially to produce evidence for the chain model. Relative error methods reveal that relative error of temporal variability is larger for smaller intervals than for larger intervals (e.g., a complete sentence). Allen (1968) and Koshevnikov and Chistovich find this significant; it appears to provide further evidence for the comb model. But as Ohala (1975) points out, the relative error values evidenced are directly derivable from the definition of relative error itself, and therefore provide no useful evidence for the existence of the chain or comb model in speech.

Abandoning the statistical methodology of Kozhevnikov and Chistovich, Ohala (1975) investigates speech timing--feedback and speech-timing perception relationships for evidence of the applicability of the comb or chain models to speech. "[The] difference between the comb model and the chain model is that in the former no sensory feedback is used to determine when a given gesture will be executed, whereas in the latter sensory feedback is used for this purpose" (1974:448). Ohala intervened in the feedback system of his subjects' speech processes with masking noise and anaesthesia of the tongue.

> Each [subject] spoke a corpus of [one hundred fifty] sentences [representing a randomized ordering of three sentence types, each type of which was spoken fifty times] in their respective languages under three experimental conditions (and in the following order): (1) control,. . . no sensory impairment; (2) masking noise, in which broadband noise of sufficient intensity to mask speech was fed to the subjects's ears over earphones, and (3) anaesthesia, in which the surface tactile sensation of the subject's tone and palate was reduced slightly by an oral application of Xyolocaine Viscous (1975:449).

Again, results are not conclusive. For the Japanese speaker variability was similar under all conditions; this provides evidence for the existence of the comb model in speech. For the English speaker variability was less for the anaesthetic condition than any other, indicating the chain model. Ohala's timing-perception discussion documents perceptual import for short-term variations in speech timing--at most adjacent syllables--but no perceptual import for long-term--e.g., phrase or sentence length--temporal variation or precision. From these findings Ohala argues for a hybrid comb-chain model governing the temporal structure of speech. He posits a "pre-preprogrammed time schedule," the comb model, for short spans of speech, "say over one or two syllables," but no time schedule, the chain model, for longer structures of speech and long-term timing.

Huggins (1975), in an investigation of the extent of the domain over which a word influences timing in a sentence, finds evidence in support of an isochronous, or at very least comb, model in speech and in contradiction to Ohala's (1975) final hypothesis. Huggins created a sentence composed of three long stressed syllables into which four unstressed syllables could be independently inserted:

Cheese(s) (a)bound(ed) (ab)out.

Example II.1

He had two speakers read each of the sixteen possible sentences five times in an irregular order. Segment durations were measured from spectograms. He hypothesized that if the word were the critical variable, "one would expect to find progressive shortening of the stressed vowel [with the addition of each unstressed syllable in the foot]" (1975:459). Huggins (1975) found a difference in effect for addition of unstressed syllables to "cheese" and "bound." The stressed syllable "bound" was shortened by the addition of unstressed syllables whether or not they were within the word, supporting the foot as critical unit, but "cheese" was shortened only by the addition of unstressed syllables within the word, contradicting the foot as critical unit. He concludes that stressed syllables "affect the duration of the preceding stressed syllable except where blocked by an intervening syntactic boundary" (1975:462). This work seems to point to the existence of metric feet in speech and thus to the existence of some form of isochrony. Huggins closes with a final suggestion: "the metric foot might be more influential in fast fluent speech than in citation form" (1975:462).

The conflicting and ambiguous evidence from this tradition would seem to speak of either serious flaws in the methodology or models employed or of some basic ambiguity in the data itself. While it is possible that speech is a fundamentally ambiguous data source, there are several problems in the models employed. The focus on language in isolation, i.e., the use of prepared texts in experimental situations presents a methodological bias that cannot but influence experimental findings. Zlatoustva (1975) found that speakers, when presented with a text form which all graphic regularity and rhythm had been removed, could readily determine the metrical quality of the passage from the sequence of regular rhythmic structures in the passage and would adapt their speech pattern to it. Mair (1978) argues that "even" a written text imposes its implicit cognitive and thus physical rhythm on the reader, when it is read. A prepared text may well introduce a confounding variable. We must study spontaneous speech if we are to determine any real temporal structure in speech.

The focus on micro-acoustic units may also incorporate difficulties into the study. As Mair (1980) has noted, the limits to measurement are the limits to technology, but it is the perceptually significant units for interactants that have import for the study of talk. They go together to form the system of human communication. And we are not at all certain of the limiting boundaries of human perception. For example, Huggins (1975) found that under some circumstances listeners were sensitive to changes in timing as small as a few milliseconds, whereas Mair and others have proposed .1-second as a perceptual threshold in interaction. They argue that changes occurring faster than .1-second tend to blur and run together into a single continuous perceptual unit, much as successive still photographs passing before the eye at a rate faster than one twenty-fourth of a second blur together forming the perceptual whole of a motion picture. Identification, definition, choice of units, and demarcation of unit boundaries is still problematic, then. We must be careful, specific, and explicit in our choices.

Perhaps Ohala (1975:451) has the final word on this whole tradition, at least for the purposes of this investigation. "We are asking how the timing of speech is regulated, but we should first ask what the perceptual value of the time structure of speech is" (emphasis added). The body of research devoted to that constitutes a tradition of major consequence.

2.3 Studies of Time and Perception in Speech

Those who have chosen to investigate the perceptual value of the temporal structure of speech (language) have generally identified the parameters of "talk" as tempo, timing, stress, and intonation. Pauses and hesitations in the speech stream are taken as evidence allowing the assumption of temporal structure in language. At least three identifiable bodies of research focus on temporal structure to address three different kinds of phenomena. One views pauses as evidence for cognitive process. A second looks at affective domains as demonstrated by pause and hesitation in speech. And a third, the most recent, investigates timing as a social phenomenon. Needless to say, these categories, although analytically distinct, often fuse and overlap in the reality. An individual researcher, a single piece of research, may trespass on all three. Chapple, in one of the most coherent of the early studies, argues for the importance of timing in "ordinary talk" as indicator of both cognitive and affective states, and a specification of turn-taking between speakers in interviews.

In addition, a fourth body of research, one which at least potentially unifies and incorporates these three, has latterly emerged. It describes a basic, perceptually significant rhythm in speech. As evidenced in the work of Erickson (1980), Mair (1977, 1980), Urion (1978), and

Chapple (1980), among others, this newly defined body of research moves beyond timing and rhythm to consideration of the totality of the process of communicative interaction. It argues that study which addresses the whole of the process of human interaction may allow us to begin to account for the often intuited connection between speech (ordinary talk) and its biological foundations in the organism which produces it, between discourse and the communicative context of which it is a part. As part of a more general move toward a "biology of communication" it may begin to allow us to get at what is really going on. It is a new perspective from which to look at talk.

2.3.1 Pauses and Timing: Methodological and Theoretical Problems

2.3.1.1 Pauses as cognitive process

Research on pause and hesitation as evidence for cognitive process assumes a stimulus-response model: the amount of time between presentation of a stimulus and the subject's response to it is taken to index the number or difficulty of cognitive processes involved in the production of a response. From this model Goldman-Eisler (1968) argues that pauses in speaking occur immediately before major lexical items. Pawley and Snyder (n.d.) extend the argument to clauses and sentences. They argue that speech is processed one clause at a time. Any more complex grammatical structures require "further processing" and thus are subject to longer pause times. In a later work, Goldman-Eisler (1980) concludes that only pauses at clause boundaries are available to listeners for cognitive processing. The listener is prepared and able to use pauses at clause boundaries because syntactic and intonation cues enable the prediction of clause boundaries and placement of pauses there. The location of other pauses in speech is held to be unpredictable for the listener.

Chafe (1980) provides corroborative evidence. He found that the length of the pause in speech was correlated with the size of the cognitive unit processed. His experiments took subjects from various cultures, exposed them to a silent film, and had them recount "the story." Scollon (1981a, 1981b) in an extension of this view, argues that in spontaneous oral narratives storytellers use pausing as an indicator of units of narrative structure. He claims: "pauses interact with sentential units to produce marking of lines, verses, stanzas, and scenes (1981a). Sabin et al. (1979) found shorter pauses in easy material such as prose readings and progressively longer pauses for the more dense, complicated material in poetry readings, spontaneous narratives and story re-tellings.

Butterworth (1980) investigates cognitive planning procedures, looking for basic units of planning. He hypothesizes that planning induces a cognitive load, and thus requires pause time in speech. From the pattern or placement of pauses in speech he aims to infer basic cognitive planning units of language. "Generally,. . .we will still try to infer from the pattern of pauses, the nature of internal representations, Plans, that the cognitive system employs, and the processes operating on those representations, Planning" (1980:157). Butterworth concludes that semantic, not syntactic units are responsible for pause times and pause time variations.

> In unrehearsed speech, speakers engage in time-consuming Planning whereby a general semantic Plan is formulated, in advance, for a chunk of speech output. Periods of Planning alternate with periods in which the Plan is executed. The Plan. . .typically comprises several clauses. . . . The Plan appears to comprise lexical Plans, which embody a semantic specification of the lexical items required to express it; and to comprise instructions for the syntactic organization of the several clauses encompassed by it. . . .Syntactic organization is the result of a syntactic organization process, which appears to be an automatic consequence or prior semantic Planning, since it causes no output delays, and is integrated into suprasentential semantic units (1980:137).

Beattie (1980) found evidence for similar conclusions in his investigation of the location and distribution of hesitations in spontaneous speech.

These studies represent an important step in the analysis of temporal structure in speech. They have moved beyond the detailed analysis of language in isolation to a more encompassing examination of speech in context, in a wider frame of time and space. The conclusions from these studies seem to point to a semantic temporal structure, dependent on meaning and context, interacting with a more or less unconscious syntactic temporal structure, which is itself the interactive product of the syntactic structural possibilities of the language and as-yet-unspecifiable ordering device, something "in the works" as it were.

2.3.1.2 Pause as affective domain

The second body of research looks at pauses in speech as evidence of affective domains. Findings from this research apparently substantiate Butterworth's claim that semantic units are responsible for pause time variations in speech. Recent studies provide a more specific breakdown of the relationship between pause and affective state. It has been found that "mild anxiety" increases the amount of speech while decreasing response and interturn pause time; "high anxiety" produces the opposite effect, increasing response and interturn pause time, and decreasing the amount of speech. Siegman (1979) argues that anxiety is a cognitive phenomenon at bottom, the result of an individual's perception of any given situation.

2.3.1.3 Pause and timing as social phenomenon

The third body of research investigates timing in speech as a social phenomenon. Researchers in this tradition contend that any perceived temporal structure in language is affected and/or effected by social phenomena. Beattie (1980) speaks of length of pause as a "socially reinforceable" variable, and argues for a "socially-symbolic" role for pause in speech. Scollon (1981a, 1981b) argues that interaction is in fact integrated through the use of pause. Scollon's findings are based on a study of traditional oral narrative genre. Butterworth (1980) posits two communicative (social) functions to pause: pause makes time available for the listener to perform cognitive processes necessary to understand the speaker; and pause may serve to disambiguate otherwise ambiguous lexical formulations. Beattie (1980) relates non-verbal to verbal timing sequences in communicative interaction. Gaze, he discovers, is organized in a coordinated system with what he calls "temporal cycles," suprasentential encoding units connected with the notion of idea. Speaker movement and gesture also evidence integration into temporal cycles. Beattie concludes that the overall structure of paralinguistic and non-verbal behaviour is largely governed by the basic cognitive processes underlying speech and that the structure can be "usefully described" with respect to the basic psychological units of language. His cautionary proviso is of importance here. Beattie indicates that gaze patterns and speaker movements are differentially evaluated. Social factors thus impinge upon what might be optimal cognitive process, affecting and effecting performance.

Several other researchers, mainly from sociolinguistics have documented the at least theoretically important inter-relationship of pause and social phenomena. Generally, their argument has been couched in terms of pause, timing, and style. Pause and rhythm are held to be central to the maintenance and integration of the event and the shared perception of the event to which the social group in a given oral tradition attends. Perkins (1979) for example, describes rhythm "matching" in the Hawaiian oral tradition as a metric for agreement among participants. Scollon (1981a, 1981b) has found the attribution of negative stereotypes, particularly negative ethnic stereotypes, to be related to perceived pause in speech. "Cold, reserve" (negative) stereotypes are attributed to longer pause times, and "warm, outgoing" (positive) stereotypes to shorter pause times. Scollon reports that other researchers have found similar attributions of stereotype on the basis of pause time, be he wonders at the cross-cultural validity of this research. In addition, Scollon (1981a) remarks upon research which argues that field dependent individuals are more susceptible to social influence than field independent ones. Productivity and response times in field dependent individuals are influenced by their conversational partners. Women in general are more field dependent than men; from this Scollon argues, somewhat tentatively, that "social category" such as sex roles and their behaviour are strongly correlated with pausing phenomena.

2.3.1.4 Limitations of "pause" studies

The notion of pause is thus central to several large bodies of research. There are serious problems with the notion, however. There is disagreement about what might constitute a pause. Scollon (1981b) flatly states there is no coherent definition of pause available in the literature. One problem has to do with the presumed discreteness of the units involved. How do we identify beginning and end points? A definition which presumes silence equal to pause has not taken into account nonverbal behaviour which, as Beattie has shown, is at least systematically coordinated with speech, if not an integral part of the speech with movement stream. Mair (1980) has shown that nonverbal action "goes with" speech and may in fact carry the semantic import of the message. The commonly accepted Goldman-Eisler definition of "pause" as a silence of 0.20 seconds or greater, determined mechanically, is patently lacking. There are too many undefendable assumptions in that definition. Despite, or perhaps because of attempts at scientific rigour, researchers have often failed to recognise the arbitrary nature of criteria used in determining basic units. (Every description embodies an interpretation of what is being observed, of course.)

Confounding variables, such as the effect of social phenomena on pause timing and distribution, have not made identification, definition and measurement of units any easier. Butterworth (1980) found that speaker perception of social role, task demands, and so forth, have an influence on the distribution of pause. Visual stimuli, e.g., reinforcement, may also affect pause distribution. At least one researcher (Butterworth) asserts the existence of a process of ontogenetic adaptation in pause distribution such that a listener or successive listeners "condition" the speaker to pause at given locations.

The choice to focus on stressed syllables represents an effort to circumvent the problem of pauses. Lehiste, Ohala, and others appeal to an identification and measurement of stressed syllables in the speech stream in order to validate their various hypotheses about the temporal structure of/in speech. However, as Butterworth points out, the identification and measurement of stressed syllables is itself an ambiguous enterprise.

> Just to measure the time between stresses with the degree of accuracy required is not yet possible, since the crucial measurement is not between word onsets and offsets . . .but between the 'Perceptual Centres' of the stressed syllables. . .and there is no way at the moment of determining this for syllables of spontaneous speech (1980:158).

Earlier research into pause and timing in speech depended upon interjudge reliability and agreement for definition, identification, and description of "units," pauses and stresses, in the speech stream. This in itself constitutes a major methodological and theoretical flaw. As Eibl-Eibesfeldt (1967:482) argues, "every describer describes what he considers significant." Even highly trained observers are biased, shaped by the very training which seeks to make of them objective, expert, supra-observers. "Not only can fifty million Frenchmen be wrong, but they are highly likely to be wrong the more they are in agreement" (Birdwhistell 1970:78). Furthermore the units chosen as significant by the observer may not be those that are significant to the interactive participants, and it is these latter units which we have set out to investigate. As Osgood admits, studies relying on inter-observer agreement may reveal more about the semantic system of the observers than anything "in the data" (Mair 1978). The only relevant data are those which have value for the participants, those units which produce observable, perceptible changes among (in) the interactants themselves. The need is for an empirical methodology to determine what the real units are. Intersubjective agreement will not do.

In recent years researchers have evidenced an increasing concern for the objective validity of their work. They have made giant steps toward objectification in the study of communicative interaction. The use of precise mechanical devices and computer analysis has become commonplace. Jaffe and Feldstein (1970), for example, employ voice-activated microphones and computer analysis in what they call a "completely automated" recording situation. Their data yield what Scollon (1981a:5) describes as "highly refined statistical measure of sound and silence in ordinary talk." From these measures Jaffe and Feldstein make inferences about the affective states and social relationships of their subjects. Unfortunately, however, objectification and mechanical precision have not eliminated the problem of unit definition and identification. It is not at all clear that what is being measured is of any major relevance to the study.

Once again Mair's remarks are to the point. Mair (1977) argues that the limits to accurate, objective measurement are the limits of technology, but what we must search out are the perceptually significant cues for interactants, the cues upon which participants depend for interaction to occur.

2.3.2 Time as Coherence Principle: The Synthetic Approach

The following sections document what I have termed a "fourth" body of research. They outline the current work of three researchers, specifically Erickson, Scollon, and

Chapple, who endeavor to get at that elusive definition we have been lacking in our study. Their work represents three attempts to account for a perceived rhythmicity in speech, considered as part of a total process of human interaction. They seem to be making tentative steps toward a resolution of the problem of temporal structure in language. At the very least, they have reformulated the problem from new perspectives. Erickson and Scollon come from a background in formal linguistics; Chapple is an anthropologist with biological training who has been extensively involved in applied anthropology and the application of interaction principles to business organisation and medicine. They bring different and varying perspectives to the study, yet they have each, independently, chosen not to ally themselves with a particular school or paradigm; rather they draw on pertinent aspects of all. They move beyond a narrow focus on temporal aspects of speech in isolation, to exploration of the "nature" of talk in context, as part of the communicative process.

2.3.2.1 Erickson and timing

Erickson (1980) makes specific claims about the "rhythmicity" of speech. He submits that "not only is talk timed [and] not only are pauses integral in negotiating turn exchange, [but] talk itself is rhythmically timed to a regular underlying metric or tempo" (in Scollon 1981b:6). From his data, Erickson is able to show that, in an oral test situation in an elementary school classroom, the teacher sets up a regular rhythm in her speech and expects that answers will be provided by the test subject, the child, in that rhythm. His data demonstrate that only responses falling on the "downbeat" established by the teacher are considered for their validity as test question responses. Responses on other beats are heard as irrelevant to the test questions. Erickson's "downbeats" occur approximately one second apart throughout his corpus, forming a regular, metric pattern.

An investigation of junior college counselling interviews also found the rhythmic pattern of the interaction, particularly that in speech, to be a major factor in the interview, often determining the success or failure of the interview. Furthermore, and most importantly, Erickson found in this study that interactants pattern their nonverbal communicative repertoire to the underlying rhythm perceived in speech by observer and participant alike.

2.3.2.2 Scollon and tempo

Scollon's (1981a, 1981b) work follows on from that of Erickson. Scollon sees in Erickson's work the insight necessary to reframe in a more useful manner the study of temporal structure in speech. He is intrigued with Erickson's claims about rhythmicity and perceives in them a way possibly to begin a coherent address to the notion of temporal structure in speech, specifically discourse, without getting caught in the messy debate over definition of "traditional" units; e.g., what might constitute a pause or stress? Scollon sets out to examine several different kinds of what he calls "talk situations." His sample consists of eighteen such talk situations, ranging from family gatherings and breakfast table talk to radio announcers and university lectures, to Groucho Marx, to Athabaskan "tradition bearers" telling traditional narratives in both Athabaskan and English.

With regard to methodology, Scollon taped each session, made a full transcription in standard--not phonetic--notation, and marked stressed syllables in the text, first from the transcript, as a check on the transcriber's normative system, and then from the audio recording. Intervals between stresses were timed with a stopwatch. Scollon then slowed down the tape of each interaction 20% and listened for a regular metric pattern of stressed syllables. He calculated beats per minute and determined measure length by counting the number of beats between highly stresed syllables. The result is what Scollon calls a "quasi-musical score."

Scollon's work presents corroborative evidence for Erickson's claim about the rhythmicity of language. Scollon found talk in all contexts to be timed to an underlying tempo, one he proposes is most easily represented by a $\frac{2}{4}$ or duple measure. He argues that this fundamental duple rhythm is the basis of a very large number of temporal variations in speech and posits syncopy (beats anticipated), hemiola (two different rhythms built upon the same underlying metre), tempo rubato (arbitrary lengthening or shortening of the beats with compensatory changes in the length of neighboring beats), and anacrusis ("pick up" syllables or unaccented syllables preceding the downbeat) as characteristic of language as well as music. He seems then to be arguing for a complex isochronous model of speech.

Scollon contends that everything in interaction is done in time to the basic underlying rhythm evident in speech. He documents speaker exits and entrances performed in time to the tempo set by preceding speakers. The new speaker may accelerate or retard the tempo to establish a new tempo once s/he has "entered," but entrances and exits almost always confirm the established tempo. He also found that "paralinguistic" and nonverbal cues--um's and ah's, clearing the throat, and so forth--are performed in a rhythmical manner which allows prediction of the succeeding tempo "as surely as a conductor's silent 'one, two' before the orchestra's entrance" (1981a:14).

Scollon isolates three factors integral to the study of rhythm in speech: tempo, silence and density. Tempo is defined as the "speed of succession of downbeats and upbeats in a measure of $\frac{2}{4}$ time" (ibid.);

silence is a beat on which speech does not occur; and density is referred to as the numbers of words per measure, or the number of words per second. That is, density corresponds to whole, half, quarter, and so on, notes in musical notation, where a higher density would be expressed by a larger number of sixteenth to sixty-fourth (etc.) notes. Density is thus a measure of the "sense" of pace in talking. Scollon argues that density is not at all related to tempo. Silence is an indicator of what Scollon terms the "focus" of the interactive situation. He argues that the more focussed the situation, the less room there is for participants to negotiate for their perception of the situation. Focus may be affected by time restrictions, the number of participants, and the introduction of communication media between participants. The amount of silence expressed in percentage indicates the degree of focus of the situation, with the least silent interaction having the most focus.

Perhaps the most important aspect of Scollon's work, both for Scollon and the present study, is his concept of tempo. Scollon submits that tempo is, first and foremost, some "thing" which is "highly negotiable" among participants and not characteristic of either individual speakers or situations. On the basis of inter-judge reliability, Scollon argues that the tempo of which he speaks is really "out there"; it is neither an artifact of analysis, nor is it situated solely in the perception of the listeners and observers. He finds that at the "boundaries of events" (1981b:10) rhythmicity is more difficult to perceive, and he attributes this to the fact that it is at these boundaries where participants negotiate for tempo. Tempo "appears to be used as the means of negotiating the interaction between speakers" (1981b:8). Tempo is the mechanism used by participants to maintain and integrate the interaction.

Scollon claims he is looking for something that will enable him to discuss the conversational equivalent of "quality" of musical performance. From the analogy with music (which he argues is more than analogy) Scollon hits upon the idea of "ensemble." "Ensemble in music refers to the extent to which the performers have achieved one mind, or to favor Sudnow (1979a, 1979b) one body in the performance of their work" (Scollon 1981b:12). Tempo is the central element to the achievement of ensemble. Tempo regulates the temporal structure, the placing of notes in time. "Tempo is the temporal bond that allows us to move together in real time. It gives us an account of the immediate past and a basis for predicting the immediate future" (Scollon 1981b:13). Scollon asserts that temporal predictability is the necessary structure, the bond, on which both musical performance and communicative interaction depend for their existence. The "temporal bond of ensemble" holds participants together in interaction, guiding them through the labyrinth of possible communicative interactions.

Scollon maintains that his work, and that of Erickson, constitute a conclusive proof of the "fact" of the rhythmicity of talk. The further postulation of a mechanism which regulates and maintains interaction--tempo in ensemble--allows us, he feels, to begin reasonably to speculate on the nature and functioning of the biological organism involved; that is, does the rhythmicity of talk derive from the actual physiological functioning of the organism itself? Scollon seems to be looking for the "reason," the mechanism, which might account for why and how it is that we, as communicative interactants "tie in" to the rhythmicity expressed in the temporal bond of ensemble. He is searching for a rhythmic, biological oscillator underlying talk.

Scollon proposes the heart. He argues that the human heart has both of the qualities necessary to qualify as the biological oscillator underlying talk: it has an independent, biological rhythm, and a susceptibility to influence from "what we might call mental states" (1981a:21). Scollon observes that the range of the rate of oscillation of the heart corresponds almost exactly to the range of tempos of talk he found in his data. Furthermore he argues that it is as a result of the "rhythmic ensemble of their talk" that participants in interaction tend to approximate one another's heart rate; and that the average heart rate of a given group in interaction closely approximates the average tempo of their interaction.

Scollon correlates quicker heart rates, warmer bodies, and faster tempos of talk (1981a). He argues for a connection between behaviour, rhythmicity and information. Rhythmicity increases with an increase in cognitive load and/or rate of information processing, yet rhythmicity is also a demonstrably and particularly social behaviour. (There is evidence which documents an increase in the rhythmic behaviour of children playing in groups versus playing alone.) Scollon concludes:

> . . .in the social context it is ordinary talk which provides the primary mechanism of entrainment of the individual with the environment and that in the individual it is the heart rate which is the underlying oscillator of this social ensemble (1981a:24).

Thus Scollon attempts to document a biological foundation for the perceived rhythmicity, the temporal structure, of talk.

2.3.2.3 Critique of Scollon

He presents an intuitively appealing case. There is a "rhythmicity" in talk; everybody knows that. There are some problems with his argument, however. First, his methodology, arbitrary definition of units, reliance upon interjudge reliability and agreement for identification, definition, and validation of units, is very simple technologically. It may also constitute a major flaw in his very interesting and intuitively appealing work. Scollon claims he chose "stress" as his basic unit of analysis because his first intuitive sense was that "the metre was being marked by the progression of stressed syllables" (1981a:9). His work then is based on judgmental evidence, validated by interobserver agreement. But as has already been noted, there are serious challenges to the validity of such an approach, particularly in the study of communicative interaction. If every describer in fact describes what he finds significant, then Scollon may be merely documenting a particular, shared, semantic system which may, or may not, correspond to that of his subjects. That is, it may or may not correspond to the system "at work" in his interactions. He simply cannot tell.

Scollon argues for the applicability of his method in this instance. He asserts that "as with music, metronomic exactness is neither achieved nor desirable in ordinary talk" (1981a:9). Tempos vary over an underlying regular metric rhythm and methodology must allow for such variability. Scollon argues that the basic problem is the question of whether the tempo itself or the perception of the tempo varies--in both participation and analysis of interaction. For Scollon, the check of interobserver reliability and agreement--that is, marking stress independently from the written transcriptions and the audio recording--in conjunction with the observation of synchronous speech and movement of the original interactants on the judged downbeats and upbeats is more than sufficient evidence to support his claim for the validity of his methodology and data.

Scollon raises some important points about the study of interaction and argues a fairly convincing case. However, his argument is limited in that he does not address, as he must, the "other"--the problem of data acquired and validated through interobserver agreement--which, in itself, constitutes a fundamental challenge to his approach and the results he has obtained.

Second, Scollon evidences some confusion in his statements about the nature of the rhythmic oscillator underlying speech. He contends that he is using the notion of tempo as a harmonic oscillation and bases his argument on the harmonic oscillator model. A harmonic oscillator depends on a steady input of energy to produce some form of periodic oscillation, usually represented by a sine wave (Scollon 1981b). But the fundamental oscillator in his model, the one which is the underlying oscillator of the social ensemble, is the human heart, by definition a relaxation oscillator of the Van der Pol type (Chapple 1980). A relaxation oscillator consists of a latent phase, in which energy is built up, and an active phase, in which energy is discharged. The "off-on" cycle repeats continuously; it is necessary because of fatigue in the biological organ or organism. Relaxation oscillators are nonpendulum, nonlinear, nonstationary. They have the "ability" to shift their values radically from one cycle to the next; that is, they do not always enjoy the dynamic equilibrium of a harmonic, pendulum-type oscillator (Chapple 1980).

The notions of tempo, ensemble, and the temporal bond of ensemble are also problematic. They are basically descriptive, not explanatory. Scollon is unable to adequately account for either our ability to be "tied in" to interaction, or our ability to opt out. He is reduced to an assertion of his description. His notion of temporal bond appeals to the concept of entrainment, but the entrainment model is at bottom a stimulus-response model. It is too simplistic and makes no provision for individual allowance of control; i.e., there is no way to talk meaningfully about the fact that an individual must allow interaction to occur. With the entrainment model we cannot account for the fact of individual conscious awareness: we cannot account for the fact that one may be immersed in an interaction, entrained as it were, and still be cognizant of oneself, even to the most mundane details such as a pain in the back or neck. Such "awareness," it may be argued, enables us to "opt out" of interaction, to interrupt or refuse entrainment, and must be accounted for in our explanation of human speech interaction.

Finally, Scollon seems to assume a descriptive adequacy for Western musical theory and notation that ethnomusicologists, among others, would envy. As Charles Seeger (1958) has carefully pointed out, the Western system of musical notation is prescriptive, not descriptive. The system is imbued with a definite and limited perspective on the nature of musical sound and it demands knowledge of a very large proportion of that perspective for comprehension by the reader. Buder (1980:32) summarizes the problem: "As a language [the Western system of musical notation] only indicates instructions for the actions, not the descriptive feeling of the action." A central problem in the theoretical monument of both ethnomusicology and musicology proper has been the creation of that which Scollon assumes, an adequate descriptive device for musical sound. Again, Buder's remarks are to the point:

> The musical continuum is not homogeneous, it is repeatable, malleable, and elastic. When the events are slow paced, time seems to pass more slowly. When the events occur quickly, time itself seems to rush headlong. What is real time in music; the homogeneous time of underlying number and measured duration of observed results, or the non-homogeneous intuitive times underlying repetition and the normative pace of performed action? (1980:30).

2.3.2.4 Chapple and entrainment

Chapple has been involved in the study of interaction and language for more than fifty years. He is concerned with establishing the biological foundations of all aspects of human behaviour--interaction, language and culture. He argues that because we are dealing with biological organisms we would be well advised to at least look at the ways in which their biological properties influence the entire range of what we customarily call "cultural phenomena" (1980:744). In a recent overview (1980), Chapple provides an excellent summary of his work.

His model is a relaxation oscillator. Chapple's claim is that biological rhythms underlie every facet of the functioning of the human organism. They must therefore be accounted for in any model which purports descriptive or analytic adequacy. Chapple submits that biological rhythms are systematic, non-random, and associative. They allow the organism to adapt to the external environment, to other organisms, and to the internal environment of internal metabolic processes. He argues that each "level" of biological oscillators, from the level of the cell on up to the complete organism and beyond, combines into progressively more inclusive wholes, each forming a "monofrequency oscillator." He continues:

> These [monofrequency oscillators] in turn are again combined. . .ultimately to become the

> action and interaction patterns of the total organism. At each level of hierarchical organisation, beats or pulse trains are the integral outputs of specific monofrequency oscillators. These create a synthesis and establish the melody (the idiosyncratic pattern structure for each person) (Ibid.).

Thus according to Chapple the patterns of and in interaction are not only organised but synthesized within each individual by all the rhythmic processes occurring in the body. Chapple calls the body rhythms the "carrier pulses on which verbal and nonverbal actions are embroidered" (1980:748). He argues that the hierarchical organisation, the "hierarchical synthesis" of oscillator populations is that which integrates the "stuff of life," sound and movement, into manageable and meaningful units of communication. He proposes that, as an integral part of the process of "living," energy is constantly taken in and converted by the organism into sound and motion thereby producing these temporal patterns.

Chapple is bothered by the persistence of the stimulus response model proposed by Pavlov, Watson, and Skinner. The S-R model assumes a universe where no response can occur until there is stimulus. There can be no spontaneous or rhythmic acts performed by any organism. The S-R sequence has been and continues to be considered the "mechanism" for description and explanation of human actions and human development. But, Chapple argues, the model is invalid; there are "whole systems" of self-initiated and self-regulated spontaneous activity that begin at the cell level, with mitotis. "Rhythms begin at the level of DNA and RNA metabolism, cell division, and develop through a hierarchical system to the whole animal" (1980:745). Biological rhythms and their manifestations through the somatic nervous system are founded in the functioning of the autonomic and endocrine system; they are all interrelated. The voluntary-involuntary dichotomy must be largely obviated; and so is the stimulus-response model.

Chapple's emphasis on the biological rhythms underlying interaction offers a unique vantage point from which to examine the process of communicative interaction. A new perspective, it obviates many old dichotomies and invalidates several old problems. Rhythmicity, for example, is included in the notion of temporal structure in speech by definition of the very nature of the human organism.

Chapple begins his work with an emphasis on the importance of the individual. "In measuring interaction, we are first of all concerned with defining the properties of each person's individuality (personality)" (1980:748). The individual, composed of the multitude of his/her biological rhythms, is the "basic unit" from which to begin the study. The biological rhythms of an individual are, he contends, built on "fixed action patterns" of the species which are shaped and selected by experiential, cultural, learning processes. Even symbolic forms take their meanings from biological rhythms.

From this basic set or frame Chapple outlines eleven "established uniformities" which make up what he refers to as "testable hypotheses" on the nature and characteristics of the process of human interaction. He posits several important principles of interaction, and delineates major implications of the realisation of the "fact" of the "complexities of combinations of biological rhythms in each personality, individualized by genetic and experiential factors to create a whole person" (1980:754). Because of their potential import to study of process of communicative interaction, I present a brief description of the eleven (as well as his "important principles" and "major implications") below:

1. Each person has his/her own rhythms of action and interaction under basal conditions, in a free-running state.
2. There is an intradian rhythm which regulates manifestation of interaction-activity rhythm during any 24 hour period.
3. The total "amount" of interaction during a 24 hour period is constant for an individual. She/he will attempt to compensate in some manner for too much or too little interaction.

4. The human organism has a "need" to achieve "optimal" or at least significantly high "levels" of synchronization (1980:751).

Synchrony, in Chapple's terms, defines an interaction where the give and take of conversational interaction takes the form of a perfect relaxation oscillator, i.e., when the speaker is in the active phase, the listener is in the latent phase. There are no overlaps, "no interruptions or hesitations or failures to respond" (1980:748), and the perfect alternation in sequence of active and latent phases is continuous. Asynchrony occurs when participants' active-latent phase adjustments do not coincide. Asynchrony, no matter how brief, produces stress, and elicits a stress reaction. Chapple argues that because the "interactional-activity" rhythm derives from the rhythms controlled by the somatic, endocrine, and autonomic subsystems of the central nervous system, a "sufficient amount" of synchrony (the amount is held to be different for each person) will activate the parasympathetic division of the central nervous system and produce feelings of "well-being, affection, love" (1980:750). He maintains we choose our network of friends and acquaintances, our environment, to optimize synchrony and feelings of well-being.

Two important principles (postulates, actually) form part of Chapple's elucidation of this hypothesis. First, there is a "stress threshold" below which the organism will not respond to a given number of instances of a particular type of stress. Second, because the patterns of interaction of individuals are transitive, "the biological oscillators of one person require adjustment to those of another (and all others with whom there are relationships)" (1980:751). Chapple borrows biological terms for the process of oscillator adjustment--entrainment--and the pattern it takes--coupling. The limits to which an individual can adjust his or her biological oscillator by appealing to the harmonic ranges of his or her fundamental rhythm dictate the degree to which coupled oscillators (individuals) achieve synchrony, congruency of frequencies and phase ratios. Interaction thus tends to follow "an irregular path" from synchrony to asynchrony (from stress to well-being). When synchrony is in fact built upon the synchronization of equivalent fundamental rhythms Chapple claims the coupling to be "endogenous and capable of being maintained over a long period of time" (1980:751). This however, is very rare; partial coupling at one of the harmonics of the oscillator is often confused for this state of "true synchrony."

5. Network changes, changes in the individuals which constitute it or changes in the "quantitative value" of their interactions, produce instabilities and stresses in the system which lead to reorganization or possibly disintegration, of the network system, given the organism's need for synchrony.
6. Changes in the fundamental rhythm of an oscillator occur progressively until a "break point" (a point of no return) is reached; an irreversible gestalt shift occurs at that point, and a new "state" is in effect. Mathematical (nonlinear differential) equations which describe coupled oscillators in a relationship include variables able to measure the direction of change in each oscillator and the oscillator system, the relationship.
7. When three or more are interacting, pair interactions will be interspersed with interactions where one acts and two respond synchronously. These latter interactions Chapple calls set or group events. In such interactions, Chapple maintains, a pacemaker individual will appear--the one who can command repeatedly the largest number of respondents--and a unidirectional array will develop in the rhythm of the set.

8. As oscillators increase their synchronisation, cultural complexity and division of labor occurs and makes possible a unidirectional array of dependent ranks The pacemaker individual can and does from time to time initiate to individuals in all levels included in the "array." But, those in a lower level may not initiate, in set events, to those at a higher level; that would break the hierarchical structure, in Chapple's terms, the polarization. Those in a lower level may, however, initiate to higher-ups in "pair" events.
9. Each interactional set takes on the properties of an oscillator. The pacemaker establishes the frequency of the oscillator, but individuals at each level of the hierarchical order have progressive phase differences from the pacemaker's, and, the system has properties of its own that are not attributable to any one person.
10. The polarization of an interactional set produces a spatial differentiation which corresponds to the phase shift differences between individuals along the frequency gradient of the set. Perceptual limits on interaction characteristics determine spatial orientations and distance. They are species-specific.
11. Organised systems, cultures, derive from these polarized "arrays."

2.3.2.5 Implications of Chapple's Model

The implications of this complex set of hypotheses for the study of human interaction; the implications following from a realisation of the complexities of the biological rhythms within each individual and in each interactional set, are enormous. Those more relevant to the present study of synchrony in human interaction, as that is defined by Mair, bear further examination.

Chapple argues that individuals are able to coordinate their rhythms to synchronize with the pacemaker individual(s). This requires very rapid control of the central nervous system and is achieved through intermodal signals which cue orientation to others and enable adjustment to accentuate parallel response. It must be remembered, however, that parallel response, what others have called synchrony, depends on underlying body rhythms. Once the beat is established it is possible to coordinate with others with only minimal cues. Thus, Chapple argues, participants may "fine tune" their rhythms to anticipate what will come next. They "project," predicting where in time the beat will come down. (It is a process similar to Mair's "projection," discussed in the following chapter.) This "happy state" depends upon compatible underlying body rhythms. At the same time, it must be remembered that a "miss" of the beat by one or the other of the participants necessitates stopping and starting over again; that is, re-tuning interactants to the beat.

In Chapple's framework, the beat establishes melody, and then other rhythms of sound and movement may appear. Chapple proposes that sound and movement are parallel, interrelated modalities, the resultants of the biological rhythms of the organism involved. When these two modalities combine the outcome is the experience of a "heightened state of reactivity" and an increased susceptibility to entrainment (1980:755). According to Chapple, this reaction is due to the nonlinearity of the oscillators. For nonlinear oscillators, coupling with the fundamental frequency peaks the frequency.

> As the intensity of the coupling increases, a narrower and narrower frequency band results, [creating] a far more stable system than its components can achieve separately [Barlow 1960]. Thus the greater the number of individuals coupled, the greater the magnitude of the resulting oscillator, because the frequencies of all the participants are maximally synchronized (Chapple 1980:755).

The whole creates a system with unique properties. Chapple argues that this "explains the crowd, group, mob phenomena of human beings en masse." Finally, Chapple is convinced that the language of music, in his terms the biological rhythms of sound and music, is the language of the central nervous system. And he proposes that the perspective gained from this realisation grants new scope, new breadth and depth, to the study of all aspects of the human organism.

2.3.2.6 Critique of Chapple

Chapple argues an impressive case. He firmly establishes the study of interaction in the biological foundations of the human organism. His model may be extended to account even for "society" as a biological phenomenon. Chapple accounts for crowd behaviour, cultural organisation, variability of entrainment, in fact a host of phenomena that we have sensed intuitively are non-random, organised, and patterned behaviour, but which, to now, have largely defied description and analysis. Furthermore, Chapple adds several important concepts to the discussion of human interaction. The claims about twenty-four-hour periods, circadian rhythms, and intradian rhythms are particularly appealing. His notion of "true synchrony" and the ability of individuals to adjust to the harmonics of their fundamental frequency in order to accommodate interaction partners is unique. With this notion Chapple accounts for two commonly perceived phenomena--adjustment, and harmonics (of the fundamental frequency of voice, primarily)--in a logical, coherent manner while keeping his explanation firmly tied to the biology of the organism. Chapple's work is especially valuable for its presentation of the notion of "transients," and "irregular path" from synchrony to asynchrony which interaction tends to follow through time. These latter notions allow us to "get a handle" on the temporal line for which we must account in an adequate description and/or analysis of human interaction.

There are some problems with Chapple's work, however. First, he has problems with definition of units. He does not adequately specify a way to define the boundaries of the events he claims to study. Chapple began his long study of interaction with the investigation of time duration and frequency of "events" but admits to having difficulties recording "them" because he was unable to establish criteria for the determination of beginning and end points for each event. That was in 1936. Apparently the problem has not yet been resolved. Second, he is not clear about a metric for his model. Units must be definable, identifiable, and measurable. Chapple's units must be measurable or the model becomes tautological: we will invent transients where we see stress. Without acceptable units, specification of criteria of comparability, a major difficulty at the best of times, becomes an impossibility. What is more, two of Chapple's central concepts, the pacemaker individual and polarization, require some precise metrics in order to be validated. Chapple does not specify a scale, however. And finally, the "testability" of several of Chapple's hypotheses is somewhat questionable.

Two fundamental problems in Chapple's model have to do with teleology and allowance of control. Chapple argues for the "need" of the human organism to establish "optimal or at least significantly high levels" of synchrony. It is here that his argument falls into the trap of teleology. He infers motive to interactants. The "need" for synchrony is not a testable hypothesis. The history of interaction study demonstrates that we cannot usefully talk about interaction in terms of an inferred goal or motive. Yet, how are we to talk meaningfully about intent, without inferring motive? Chapple has no ready answer. He posit's a "need" for synchrony in human organisms.

With regard to control, Chapple and Scollon have similar difficulties. In spite of Chapple's initial insistence on the individual, there is no way to account for individual allowance or denial of control in his schema. Like Scollon, Chapple seems to talk in terms reminiscent of an entrainment model. As noted earlier, this is basically a stimulus-response model, and is too simplistic to account for the complexities of the "fact" of human interaction. There is no provision in the stimulus-response sequence for individual consciousness or awareness. Nor, as Chapple himself notes, is there room in the S-R model for self-initiated and self-regulated spontaneous activity, action initiated by the organism without external stimuli. The S-R model asserts a description; it does not explain. And the notion of entrainment, even if used in a biological context such as the one suggested by Chapple seems to lead to a notion of a one-way dominance and potentially ominous control in interaction. Although the fundamental notion of control is admittedly useful in that it allows for postulation of such potentially useful concepts as the pacemaker individual and polarization, to consider control in this manner (i.e., as a one-way dominance phenomenon) prohibits consideration of other central aspects of the process of communicative interaction. The most significant of these, individual allowance of control, has already been mentioned in the critique of Scollon's work. The entrainment model may also account for Chapple's failure to consider imperfections in the communicative process. He makes no attempt (indeed he makes no mention of) miscommunication or misunderstanding, two of the most common communicative "results."

Finally, the basic S-R quality of the entrainment model might underlie Chapple's choice of the level of description. Chapple maintains his description at the level of individual features; that is, he describes attributes of actors and features of the "controller." Yet his own data, in fact his own model, indicate the imperative to move description and analysis to the level of systemic properties if our description and analysis are to be adequate to the perceived reality. In an adequate description of the coupling of two oscillators, for example, it is necessary to address the system that includes the two oscillators as well as the oscillators themselves. There is a mathematics which describes the behaviour of coupled oscillators (of both harmonic [pendulum] and relaxation [Van der Pol] types). Chapple himself makes mention of it. It is a very different mathematics from that which describes the behaviour of single oscillators. Description and analysis must thus be contextualized in terms of the surrounding environment.

2.3.3 Synchrony

The most elaborate claim based on synchrony is Mair's. His work is summarized in the following chapter. Mair's is a unique construct, one which draws from several and various traditions in the study of human interaction.

Chapter 3

Summary of Mair's Model

3.1 Faces and Gestalt Transforms

Mair began his study focused on the face:

> ...it was thought that the study would more quickly expose the heavier methodological issues involved in human behaviour, while necessarily retaining an anchor to the physical fact of faces and the factuality of their movements (1977:IV-5).

Facial movements are recognisable, observable and quantifiable behaviours of human interaction. Mair notes that "emotion" words are the terms most commonly employed in traditional studies of facial interaction. However, as noted above, Mair argues that words for feelings demarcate sectors of experience that owe their identity to the cultural scheme of which they are a part. Thus the appeal to "emotion" words constitutes a major flaw in the traditional studies of facial interaction.

Mair argues further that words are by definition inadequate to capture experience: words chunk, apparently rather arbitrarily, the continuuum of reality into discrete categories. Words attempt description, but complete description is replication. Mair suggests that the inadequacy of words to experience may serve to account for the problem of paralinguistic researchers as they investigate human interaction:

> [P]recisely as researchers, both psychologists and linguists, try to operationalise 'feeling words' selected to demarcate sectors of experience, they lose the adequacy of their chosen markers to lived

experience" (1978:29).

Such categories, feeling words and "lived experience," are of different domains and translation is a dubious undertaking. We must attend to the applicability of concept and category to our concern.

> There is evidently something pan-human about facial movements, but if we think about that something as Emotion, we necessarily bring with the concept a lot of cultural baggage that although real enough...might more appropriately be considered as the subject of a good question about human behaviour than an answer (1977:IV-8).

For these reasons, working from this perspective, a decision was taken to avoid a number of approaches to the study of facial (and human) interaction. Mair summarizes the approaches he has chosen to avoid in his study. First, the approach which matches behavioural variables to neurological structures or genes is questioned: "it is the concept of 'behavioural variable' which is itself being questioned" (1977:IV-3). Second, Mair wishes to avoid the approach which seeks to match up behavioural variables to natural science measurements of physiological states. Third, Mair questions the worth of the approach which attempts to validate behavioural variables by statistical techniques in populations; it is the "mystification" of statistics (used here in the sense intended by Marcuse), the process of couching notions in extremely complex representations, so that only the most expert and determined search may detect possible errors) which Mair questions.

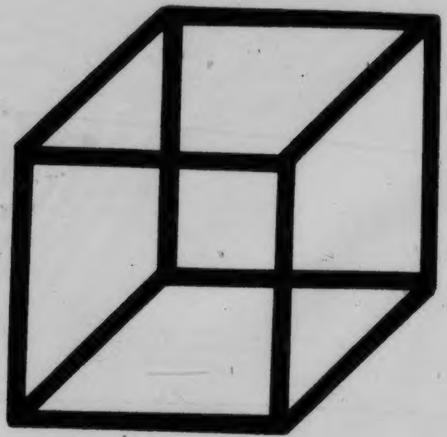
From his study of facial interaction, Mair argues that perception of the face is "categorical," occurring in the form of "gestalts." He uses the notion of "gestalt transform" (GT):

> The idea is that a particular configuration of the face would be perceived as constant while the proportions of constituent components changed, until a breakpoint was achieved, when the display would be perceived as a different Gestalt. The idea, then, of local stability to perception of a system, with sudden discontinuities, in that perception was compatible with what happens to us when we perceive a Necker cube--it's either one way or the other, but not both (1977:IV-12).

The paradigm for Mair's claim for categorial perception of the face is, then, the Necker cube. In perception of the Necker Cube there is an untimeable "flip"--not a transition--between stable states. The "front" surface of the cube "flips" from "front" to "back" surfaces, but we cannot trace its journey nor can we suspend it in midflight. Mair argues for a similar situation in our perception of facial configurations. We witness the change from one facial gestalt to another, but not the changing. Mair's claims of categorical perception of facial configurations follows on from the work of several other researchers--Mair mentions specifically Lieberman (1967) and Condon and Ogston (1966)--who have developed in whole or in part similar notions of gestalt perception in human interaction.

An important feature of the GT model has to do with its existence in time (after Bullard 1980). With the Necker cube model, the image is static and we flip back and forth between the two gestalts. This is not the case in visual

Figure 3.1 The Necker Cube



perception of non-static forms such as the face, however. When the breakpoint in perception of the face (or any non-static form) is reached, the old gestalt has all but disappeared. The interactant must rely on memory or external data, a videotape for example, to reconstruct the old gestalt. Due to the time factor, the interactant is almost literally forced into perceptual awareness of the new form. Once the new gestalt has been achieved in time, attainment of the old gestalt is not a matter of imperceptible shifts but a "catastrophic" (in the meaning of the French cognate) leap into the past. That is, the shift of facial gestalts, the shift of non-static visual forms, is unidirectional through and with time.

There were several problems with the GT model at this point in Mair's investigation. A major problem revolved around identification and definition of units. That is, by what metric are we to divide the fact into "units" for analysis? How are we to be sure that the units chosen for analysis are in fact the ones the brain is using for facial perception during interaction? The difficulties encountered in definition of the seemingly obvious unit, "mouth," offer a typical example:

> Attempts made to trace out the outline of the mouth from the film quickly demonstrated that the drawing of the line required decision, i.e., was not self-evident. And when followed frame-by-frame, the mouth appeared to move in a kind of swirling indeterminacy of shape, such that it was quite impossible to select one momentary configuration as being a definitely achieved position, rather than another one. Yet, minute changes in the configuration of the mouth quite transformed the perception of the face (1977:IV-13,14).

Definition of an appropriate time scale for the study was also problematic. Finally, there was the problem of speech: what, if anything, was to be done with "talk"?

Mair's solution for all but the final problem, speech, involves implementation of two separate procedures in his analysis. First, he argues, we must identify the "breakpoint" between perceptual facial gestalts. This is an admittedly subjective process, insufficient in and of itself to satisfy the requirements of a scientific investigation. The second procedure is an alternative:

> An alternative to the breakpoint idea--subject oriented--was that by studying the movements of the face of the other participant in an interaction one might determine which had been the crucial changes which changed one perception into another, and then they would be the movements (1977:IV-14).

He hypothesizes that facial movements in one participant cue or "trigger" facial movements in the other participant. Investigation must focus on interaction, exploring a "social physiology," in Mair's terms, that might illuminate aspects of what he has called the "social brain."

3.2 Speech Parameters, Visions, and Plans

Mair recorded several conversations on audio and video tape, and after several hundred hours of intensive observation and analysis came to the realisation that facial movement seems to "go with" speech in both behavioural (movement) and semantic (speech) terms. He found on his tapes "synchronies of all sorts, shared movements, intricate coordinations of timings and speakings" (1977:IV-15). Faces alone could not account for these data. Mair suggests an explanation for the source of the shared movements, the shared structurings of the conversations, lies in what he terms a "cognitive rhythm" of speech and movement to which all interactants are party. The structuring of movements appears to be "in" the arrangement of ideas--what Mair calls the "mutual elaboration of cognitive models during interaction" (1977:IV-15). It was at this point that Mair decided he had no recourse but to expand his study to include speech with movement.

> If all questions are incomplete models, and all answers are bits added to a model so as to complete or partially complete it, and this is what is happening by mutual elaboration in interaction, then it is the models which are the units, and we might expect a unity of shape in the models to be discovered in the timing, intonation contour, and movement involved in their production (1977:IV-18).

Mair's data are audio and video taped conversations of approximately one half hour length. Subjects were chosen who knew one another. They were kept naive as to the purpose of the study; they were simply asked to chat. A good deal of Mair's argument is in fact based on intensive, exhaustive study of a particular half-hour long conversation between two college freshmen, a male and a female. Of this bit of research Mair says:

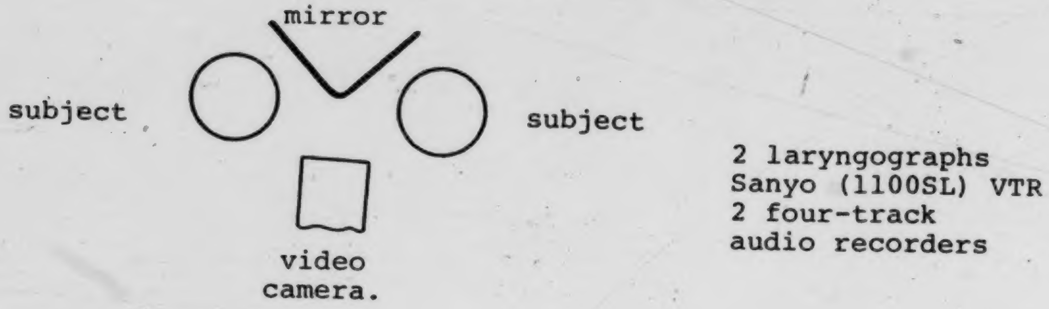
> Such spontaneity that was achieved was facilitated by their membership in the same freshman group, by the provision of wine, and by the open participation of the experimenter in the interaction. It was a sort of party. (1978:67).

Mair argues that the "control" in his experiments was "uncontrol." Nothing on the tape is not relevant data.

Candid camera techniques were forsaken for the acquisition of precise data. Subjects "chatted" sitting, with a convex, angled mirror system between them, around which they had to peer; encumbered with presumably somewhat uncomfortable laryngograph impedance detectors attached to their throats, and sporting black dots on their faces. Various recording devices, eventually a video camera, two laryngographs, a Sanyo VTR, and two four-track audio recorders occupied the room with them. This is shown diagrammatically in Figure 3.2.

The laryngograph provides measurement of basal intonation contour, voice onset, and offset, and total speech pressure wave. The video recording allows manual frame-by-frame plotting of head movements, monitoring a

Figure 3.2
Mair's Recording Set-Up


mirror
subject
subject
video camera.
2 laryngographs
Sanyo (1100SL) VTR
2 four-track audio recorders

single point of the face (a dot, or the medial canthus of the eye), and later, a computerized display of three-dimensional nose movement plots ("x", "y", and "time"). A verbal transcription in phonetic and standard notation was made and placed "in time" where the utterance occurred. The result is a multi-faceted transcription detailing simultaneously fundamental frequency contours, timing of vocalisation, and head movement "within and between" interactants.

The data are evaluated, "considered" in Mair's words, in terms of several "artificially extracted parameters": intonation/fundamental frequency contour, rhythmicity, supra-glottal modification, non-articular movement, and synchrony. With these parameters Mair proposes parsing the complex interweavings of sound and movement in interaction. As a way into the theory proper, then, the following discussion focuses upon Mair's "artificially extracted parameters."

3.2.1 Fundamental Frequency Contours

Mair posits two important features of fundamental frequency contours. First, the term "trajectory" is invoked. Mair found that laryngograph recordings of shapes outlined by fundamental frequency of voice "were indeed trajectories" (1978:27). That is, intonation contours take on a regular, discernable shape in conversation; which, once begun are committed to self reference for completion. The basic shape is analogous to that of a sine wave. It may be characterized by rising and falling "S" shapes, though the shape of a sine wave is much more nearly regular. Mair found that intonation contours appear to continue "across silences," to be taken up again by one or the other of the interactants at the point to which the intonation contours would have gotten had they been continued through the silence:

> Observation of the fundamental frequency contour for numerous utterances demonstrates smooth shapes, which continue across pauses and silences to be continued where they would have got to if the pause had not occurred. We propose the term 'trajectory' to express the slower wholeness of tune over utterance (there is also a finer embroidery corresponding to segmentation). It is the shapes themselves that suggest they be considered as trajectories (1977:V-6).

Like the trajectory of a ball, the initial form implies the rest of the fundamental frequency contour; it is in this way that the path of the intonation contour through time may be said to describe a trajectory. Mair notes that another researcher, O'Connor (1961), has made claims about fundamental frequency contours. O'Connor proposes a finite number of tunes (intonation contours) in colloquial English. He argues for ten of them, and says that these ten are apparently universal.

As a second feature of fundamental frequency contours, Mair contends that articulation, which he describes as "minute embroidery" on the general figuration of the intonation contour, "patterns" the overall shape of the wave form (1978). That is, he argues that articulation, which in laryngograph traces appears as bumps on the "carrier wave" of the fundamental frequency contour, serves to shape the basal contour, directing it through the rising and falling S-shapes by which it is characterized.

3.2.2 Rhythmicity and Plans

Rhythmicity data seem to argue for a regular, though not perfectly metric, division of utterance into rhythmic pulses. Mair maintains that the pulses in speech are rhythmic and are clearly discernable by observer and interactant alike. Mair proposes that it is the shape of the intonation contour, the rising and falling S-shapes which divides utterance into rhythmic pulses. "A trajectory occurring in time divides up time just as a pendulum does (the swing of a pendulum is a sort of trajectory)" (1978:17). S-shaped falls create rhythm, imposing on time.

Mair argues that there is no "natural" unit of time; that all units of time are in fact created:

> There is no natural unit of time; all units of time are created by plans in action, whether those plans be the physical laws which inhabit a swinging pendulum, or the rhythms of speech. Plans impose on time (1978:25).

Mair defines both utterance and action as "plans in action." He calls these plans "states of affairs" (1977:VII-3). Utterance and action are held to be physical manifestations of plans formulated in the brain. "Because utterance is a motor plan in action, the plan for it must be 'in the brain' prior to its commencement (after Goldman-Eisler, pauses within utterances can be interpreted as 'think points')" (1978:24). A complete utterance is a plan that has been made manifest in time. Mair holds that plans are beamed out onto the world through utterance and action in the same way as the plan for a tool is held to be projected out from the brain through the eyes onto the physical forms which are transformed into "tool." Thus plans in action segment interaction, creating the rhythms and melodies by which it is characterized.

Plans, specifically the trajectories of their manifestation in time, impose on time, creating rhythms, units, patterns in the world, as words are held to impose on reality, chunking the continuum of the world into discrete categories. Plans themselves are argued to be timeless constructions in the brain. Mair maintains that it is the time taken to enact plans which divides up time, and that in utterance,

> this is literally the rhythm, because rhythm is determined by certain pitch variations (particularly S-shaped falls), and these carry the semantic 'point' of the utterance. The point of the plan is its achievement, the point of utterance is the S-shaped fall (1977:VII-14).

The trajectories of Mair's fundamental frequency tracings are the paths by which plans are achieved or perceived.

Mair (1977:VII-15) notes a basic difference between dramatic performance and conversation with regard to "plan": in the former the plan is worked out in advance, and in the latter the plan is in evolution during the encounter. In conversation, interactants are literally "making it up" as they go along. Mair submits that any "immutability" in the plan of conversation is thus found not in the conversation itself but in memory, or in an external record (audio or video tape for example) of interaction. "Although plans may project the future, they act on the present" (1978:70). The ethnomethodologist's concept of "logic in use" is, Mair suggests, particularly apt.

In a summary statement Mair clarifies, opposing the concepts of "plan" and "description": "A plan creates categories, which are then truly there in the reality. A description seeks to identify categories there in some other reality than itself" (1977:VII-19). In the case of physical categories such as buildings, tools, and so on, this may "quite obviously" be seen to be the case, but Mair argues for an identical situation in the "virtual" world of language.[5] Mair contends, as noted above, that time itself is segmented by the time taken to enact, i.e., to manifest plans: "the plans themselves, being manifested in other people and the environment. . .segment time" (1977:VII-11).

Mair describes human reality as in fact consisting of "stable states of achieved or perceived plans, and the paths by which they are achieved or perceived" (1977:VII-4). He posits a unique difference between human and animal reality:

> Although motor actions performed by apes and other animals must necessarily be informed by plans, even to the fashioning of rudimentary tools, it would seem that the human characteristic is to take over time completely and live immersed in a constructed universe (1977:VII-15).

He contends that humans live in a world entirely composed of plans. The unique human extension or application of the "time trick"--that is, the ability to spend present time organising for another time, creating plans in the present for future use, permits and perhaps requires us to live immersed in our "constructed worlds." Our everyday world is, Mair argues, a "mask and support for a certain kind of society" (after Mary Douglas [1975:247]). Mair uses the example of tool construction. Tools, he maintains, are constructed in a "now" to be used in a future "now."

Furthermore, in constructing a tool the eyes "project the plan in the brain onto the physical medium" (1977:VIII-19). Tools and artifacts are manifested plans. "A tool and its uses form a conceptual unity, and each can be considered as a frozen cultural memory" (1977:VII-6). If tools can be seen to freeze cultural memory, then the use of tools demands that the user at least acquiesce to the notions of reality of which the tool is part and product. Mair argues that, in the end, "the delineation of a task and allocation of time for it, are parts of the cultural conception of reality" (1977:VII-6).

---

[5] The term "virtual" Mair takes from the term "virtual image" as it is used in optics.

The takeover of time is proposed as the "chief" of the neurological capabilities of the human brain (1977:VII-11). Mair submits that it is the brain which teaches the eye to see: "what it sees is then written into the brain" (1977:VII-11). Mair refers to the work of R. Gunter and R.L. Gregory, among others, which support his notions of the workings of brain and eye. Gregory (1970) puts forward a theory concerning the structure of language: he submits that the structure may derive from the way animals structure the world in order to see it. R. Thom and R. Gunter argue that, apparently of necessity, the brain separates the continuum of reality in segments, patterns and time. The brain "makes categories out of continuities, it 'isolates out' discrete entities in what would otherwise have been a continuum" (Mair 1980:4). The individual then may be considered as both an organising centre and an organised centre of the world. "It is as if the brain 'sucks up' the world at a certain age, and then is for ever more, an organising centre of that world" (1977:VII-11). Thus the takeover of time by man assumes cardinal import to an investigation whose intent is to uncover the principles of text regulation which govern the process of communicative interaction.

Mair argues that we must alter our perspective; we must account for the fact that "human reality appears to come in quanta or chunks, and each chunk is an action performed on the constructed world of the culture, a re-organisation or disorganisation of it" (1977:VII-17). "Action segments time. It also delivers a world that has been conceptually rearranged" (1980:10). All categories are ultimately cultural, then, and individual ways of making sense of the world depend on and are possible only in a particular culture.

> The perception of temporality is built into the structures (cultural) through which the individual awareness makes its path. We have so thoroughly taken over time that we can have no conception of it other than that of the cultural universe within which we are immersed (1977:VII-5).

Moreover, Mair argues, there must be agreement of the nature and form of the constructed reality for individuals and cultures to survive. Individual perceptions must be largely in accord or the system and the individuals which form it function improperly and cease to exist. Thus, Mair argues, reality itself--including and especially the constructed reality in which man immerses himself by virtue of his consummate ability to perform the time trick--is uniquely and fundamentally rhythmic, segmented by the manifested plans which form it.

3.2.3 Rhythmicity and the Single Central Patterning Program In the Brain

Rhythmicity data also argue for the existence of single central patterning program in the brain which structures both speech and movement. In his investigation Mair plotted out movement as a velocity versus time function, and found that peak movement velocity occurs at, or nearly at, the place of vocalic nuclei in the utterance.

> The nucleus of utterance coincides with accent and pitch change, and is always semantically important . . . It is as if the intonation contour is always just appropriate to where the model has got to in its elaboration (1977:VII-7,8).

The time scale involved is very small, one tenth of a second or less, and leaves no time for a feedback loop between speech and movement patternings. Furthermore, movements that accompany utterance and pitch change seem to have an onomatopoeic character. Movement seems to be "going with" utterance. The visual (movement) and aural (speech) seem to be patterned by shapes manifesting in both (1978:18). Mair applies the term "transmodal," maintaining that speech and movement are "turned out," structured, by the same mechanism, the same patterning program in the brain.

> Seeing pitch as squeeze [i.e., as controlled tonal variation in vocal production], and elaborating on its corollary that it is motor, and that its perception as melody is an accident of modality, then it becomes easier to accept another proposition, that speech and movement are turned out by one and the same patterning programme (1977:VII-16).

Mair holds that little movements that go with pitch change and utterance are part of the evolving model; they help deliver the semantic point of utterance. The little movements are not themselves "signals" of anything. Mair argues they may not even be construed as "emblems" or "illustrators" as Ekman describes them. They do, however, trigger shared realisations, awarenesses. This notion fits nicely with Mair's suggestion of a "cognitive rhythm" of speech and movement as the source for the shared structurings of and in

> Little movements that go with pitch change and utterance often seem onomatopoeic. In that way they are a partial manifestation of the model in elaboration. It is all happening far too quickly for the movements to be an illustration of the ideas, and our formulation of the unity of speech and movement sees all as contributing to a state of play which is at once a topic and a social relationship (1977:VII-17).

Mair makes mention of the work of other researchers who have investigated the patterning of speech and movement. Kendon's (1967) research supports Mair's notion of a single central patterning program. In fact, Kendon specifically suggests that the same mechanism which structures speech structures non-articular movement. Bouissac (1976:163) achieves a similar perspective in his semiotic analysis of circus performance. He maintains that we must consider "trans media" messages and message formulation:

> Multi-media or multi-linear communication must not be taken to mean parallel messages (this would satisfy the definition of noise) but trans media messages. . .Meaningful trans media syntagms are produced through co-articulation and immediate succession of occurrences.

Finally, G. Leisman and J. Schwartz, and R. Held, among others, have documented coordination of saccadic eye movement with the rhythm of alpha brain waves. Their research provides a groundwork, forming the base from which Mair proposes that

> the coordination of saccadic eye movement with brain alpha rhythm might suggest a quantal 'sampling' of the visual world of which the quanta-like division of auditory text into states of play might be just a more leisurely timed extension of the same principle (1977:IX-11).

3.2.4 Vision

Mair posits a fundamental similarity of visual and auditory modes. Both are characterized by abrupt transitions (gestalt shifts) between stable states. There is local stability to perception of a system, with sudden discontinuities in perception. The paradigm for it is from the visual mode and the Necker cube. From independent physiological and psychological evidence, Lieberman (1967) documents similar categorical shifts in aural perception. He found for example, that the word say, when repeated continuously, flips from "say" to "ace" in Necker cube fashion. Mair expands on the basic concept of his Gestalt Transform model of facial interaction. He argues that in communicative interaction a particular configuration--which may be visual or aural or contain elements of both--is perceived as constant, while proportions of constituent components change, until a breakpoint is reached and a new gestalt perceived. Mair maintains that evidence for categorical perception in the visual and aural modes supports a notion of single central patterning program in the brain for the speech with movement stream. He suggests the analogy of a differential gear to account for the ways in which the central patterning program is variably manifested in the several and various sensory modalities and emphasizes that the shapes of movements are not constant across melody shapes. Speech and movement are contextually determined, integral parts of the state of play being delivered with their occurrence.

> [F]acial movement, blinking the making and breaking of eye contact [and so on] all seem to fit together as one coordinated process. I propose the analogy of a differential gear to convey how it might be that one patterning program might manifest so variably between the effector organs of speech and body movement. Nor are the shapes of movement constant across the same melody shapes. Sound is two-dimensional, and yet we have suggested that the time forms of utterance be considered as actions on states of play, their effects quite specific in context, but relative across context. The onomatopoeic character of movement. . .might be accounted for it we posit that they are partial manifestations, like intonation, of the states of play that they help deliver (1977:IX-10).

In sum, the categories of speech and movement may be "separated out" by instrumental analysis, but they do not appear to be separated out in (or by) the brains of interactants themselves.

3.2.5 Supra-glottal Modification

Mair does not spend a great deal of time in specific discussion of his articulation data. He notes that the whole area of supra-glottal modification is very controversial, and remarks upon some of the problems addressed in the literature. Most notable, he feels, are the arguments as to the "reality" of the phoneme and the various extant theories competing to refine or abolish the whole notion. However, other than to assert the obvious--that a "textbook list of phonemes is not sequentially tacked together to produce utterance" (1978:17)--Mair's primary comment of concern to the present study has to do with his notion of a single central patterning program in the brain as generating source for the speech with movement stream during interaction. Demonstrating how non-articular movement fits in with articulation could, he maintains, well add a new dimension to the debate. He refers to the work of some other researchers who have approached this "new dimension" in their work. Condon and Ogston (1966) claim that micro-movements "go with" speech to a detail as fine as the phoneme. They have documented a "natural segmentation" of movement that is coincidental with and corresponds to the occurrence of phonemes. Mair himself has found that "micro movements do certainly occur to a detail equal to the change points in articulation" and further that "intonation change can occur with this detail as well" (1977:V-8). At the same time, Mair notes, Lieberman (1972) demonstrates that sequential perception of phonemes is quite simply, beyond the discriminatory capacity of the human ear. And Ivmey (cited in Mair 1977) describes "dramatic short-cuts" taken in articulation during spontaneous speech; short-cuts which Mair confirms in his own research. Thus, Mair argues, supra-glottal modification, non-articular movement, and synchronous movement cannot be adequately approached entirely separate from each other, nor can they be considered theoretically separate from all the other categories of his model.

With respect to data, Mair observes that non-articular movement occurs at the "same" time as, in the same detail as, and with the same degree of precision as articulatory movement. It appears the "levels" of movement are all turned out by the same patterning program in the brain. Such co-patterning of articulatory and non-articulatory muscles obviates much of the debate over division of language into segmentals and supra-segmentals: they are all turned out by the same patterning program. Mair also documents gross changes in direction "at or near changes in velocity, and at articulatory transitions" (1978:18). Movement, he contends, cannot be adequately described as something that merely emphasizes or illustrates what is being said. Nor would a description of changes in direction as "relating to" the semantic import of what is said be enough. From his analysis of dyadic interaction Mair argues that cognitive change points are most effectively viewed as physical events whose occurrence is related to the rhythm and timing of utterance. He is very emphatic:

> [A]ttempts to find separate functions for the different aspects [i.e., the instrumentally and analytically isolable categories] of communicative activity does violence to the reality. . .each is part of the other (1978:18).

He concludes:

> It is suggested that the speech melodies are not "signs of" any external variables, but actions on the cognitive models at the time of their occurrence. . .In the same way that it it is proposed that speech melody is not a sign, and that to search for labels whether of attitude or emotion for particular shapes is based on misunderstanding of the process of text generation, so also might the little movements accompanying speech be considered 'actions'. The precision of their integration with utterance suggest co-patterning, their shapes are often onomatopoeic, but rather than 'signing' a lable, they seem to go with the full semantic import of the state of play being delivered when they occur (1978:71).

3.2.6 Synchrony

Synchrony, the final parameter, is a crucial notion; from it Mair has derived much of his conceptual frame. Synchrony is defined as the synchronous co-patterning of a communicative event among interactants. Mair proposes that a natural spontaneous conversation may be best considered as a single system in evolution within which interactants may be described not as separate, individual "atoms," but as nodes of the system (integral parts, organising and organised centres) of the whole in evolution.

> This supra-individual model of the elaboration of cognitive process among interactants considered as organising centres rather than atoms assumes a communion between them such that the auditory and visual channels of connection are as real as nerves (1977:VII-15).

Mair argues for a model of interaction based on "mutual involvement" in a "shared awareness." "Such mutual involvement," Mair proposes, "becomes visible when it is absent" (1977:VI-6).

For Mair, synchrony is a manifestation of participant involvement and immersion in the evolution of the supra-individual system which constitutes interaction, which is reality for interactants at the time of its occurrence. It is a unique construct, one which differs markedly from the several forms of "synchrony" found in the literature. Mair's synchrony, with its assertion of "mutual involvement" in a shared awareness" may seem to ephemeral a concept for scientific study, but Mair argues that the fact that individuals are interconnected by sensory modalities rather than by actual nerves has obscured for us the fact that communicative interaction defines and manifests a supra-individual system in evolution.

Mair once again refers to the work on categorial perception, offering the findings in that literature as support for his notion of interaction as supra-individual system. Lieberman argues that speakers of any given language appear to be "differentially sensitive" to certain acoustic shapes. Condon and Ogston, and Trevarthen have shown that speech rhythms are perceived even by infants and serve as a basis for the timing of their motor actions and auditory responses. Condon and Ogston (1966) speak of speech rhythms mirrored in the movements of very young children, and provide 96-frame per second film documentation of one-month old babies moving "in time" with the rhythms of the speech surrounding them. Larger and longer movements were found to coincide with the melody of vowel sound, and short, abrupt movements went with the staccato of consonants. Humans appear to by "synched" to the rhythms of the system of acoustic shapes of language, the melody, at a very early age. Humans are nodes, integral parts of communicative systems in evolution, according to this summation of Mair's and "the individual-as-atom conception is being negated by the data" (1977:VII-15).

Mair posits two types of synchrony. The first occurs within a single individual and manifests itself in the form of synchronous realisation of auditory and visual phenomena (speech and movement). This type of synchrony is closely related to and provides evidence for Mair's notion of a single central patterning program in the brain as generating source for sound and movement within each interactant. The second type of synchrony occurs between and among interactants, and also takes the form of synchronous manifestation of auditory and visual phenomena. This form of synchrony is closely allied with Mair's claims for "shared awareness" in human interaction, a "mutual involvement" of participants in the evolution of the single system of which they are integral parts, or nodes. He argues that "the sequencing of utterances among interactants. . .involves such close participation in the temporal structuring of the utterance of the others present that this too is a manifestation of processes occurring synchronously in the brains of those involved" (1977:VI-6).

Mair proposes that observation of synchronous co-patterning of sound and movement within interactants, and the notion that they come from the same patterning program in the brain leads very nicely, almost inevitably, to the concept of "shared awareness" among interactants. The similarity of the two types of synchrony phenomena in fact underlies Mair's hypothesis of shared awareness.

3.3 The Shared Text and Melody

3.3.1 Introduction

He begins with a notion so obvious as to be invisible: interactants are present at the same time. They share a time slice. From his videotapes of dyadic interaction--the tape of the two sisters who lived together and "seemed forever immersed in an intimate dialogue" is of particular note--Mair identifies "synchronies of all sorts, shared movements, intricate coordinations of timings and and speaking [and so on]" (1977:IV-15). Here is Mair's "social brain." Both the time frame and the close participation of interactants suggest a shared model in the brains of communicative participants. However, because synchrony is apparently related to trajectory of utterance, particularly S-shaped falls, and because trajectories are predictable, it might be argued that some sort of awareness, conscious or unconscious, of fundamental frequency contours and the trajectories they describe could account for at least some of the synchronies of movement and/or utterance evident in interaction. Something is lacking in that account: neither the very small time frame involved--one tenth of a second or so--nor the fine detail of synchrony are explicable by it. The shape, timing, and rhythm of interaction suggest an interlace concept, i.e. a shared awareness among the brains involved. They seem to be acting directly on one another, one brain immediately affecting the other, to the point where the "plan" of one becomes the "percept" of the other, is incorporated in the other's projective plan of topic, gets redirected at the one as "plan," becomes the other's "percept". . .and so on *ad infinitum* until the interaction is broken off.

> Interactants share a time slice. They are present at the same time. But detailed analysis of the timing of the physical and hence biological movements of some interactants indicates that the cognitive models are *usually* for them grasped before the utterance is complete. Furthermore, they complete each other's models. Furthermore, they come in together at the same time after silence. Furthermore, the melody of one carries on from the melody of the other, such that the melody of the topic is supra-individual. Both brains are being co-organised by *what is happening*. And the linearity of awareness in these entwined encounters, if we postulate that linearity *is* in fact the trajectory of speech with movement, sugggests that the awareness is in the text. The text is a very detailed inter-penetration of two awarenesses organising a shared reality, and the resulting shared text *was* their reality at that time (1977:VII-14,15).

Mair thus posits a supra-individual system of which the two interactants are nodes, organising and organised centres. The melody of the text evidenced in the contour of fundamental frequency traces is an integral part of the system. Melody is also supra-individual. Melody is shared. Melody is more than a "something" which accompanies the logico-semantic system in the delivery of the "message." It is not just a "sign" of action of the cognitive model in the brains of the interactants. Mair argues that the shape of the melody is action; the shape of the melody "delivers" the state of play of which it is a part. He submits that voices and movements are "hard wired" into the brain. Hard-wiring mediated by the senses, interconnecting individuals, constitutes the "social physiology" of the human organism:

> I suggest that like voices, facial patterns hard-wire the consciousnesses of the interactants, controlling the direction of discourse (1977:X-16).
>
> . . .
>
> I have termed such 'hard wiring' mediated by the senses 'social physiology'. . .and the paradigm for it is that of touch where the actual tactile patterning of the sense organs at the point of contact necessarily has the same time form for both parties. This simultaneous co-patterning of the senses of both by the action of one carries its own intrinsic predictability, in the trajectory of the time form of the caress, and the ensuing action of the other carries on from the state induced in both by the action of the first. Such action is itself not a sign anymore than an action performed on the external world delivers its consequences, but is not a sign of the consequences (1977:IX-3,4).

Mair proposes that the shape of the melody--the rising and falling S-shapes of the intonation contour, the "melody of the text--in fact constitutes the mechanism which serves to guide topics to their conclusion. "In brief, it is postulated that the shapes of melodies act on topics to guide them to their outcomes, a sort of melodic competition" (1977:IX-9).

Melody is supra-individual, something to which interactants may and must appeal in the process of communication. Speech melody, Mair maintains, performs "non-arbitrary" actions on cognitive models in the brain. He argues that in conversation interactants are constantly transforming the immediate awareness of the other, through the action of melody: "speech melody, its pitch and timing, directly wires in the awareness of one interactant to that of the other, transforming it" (1977:VIII-4). Change points in cognition are held to occur as physical events related to the timing, i.e. the rhythm established by trajectory, of utterance, particularly under conditions of "rapport" (1978). The physical trajectory of utterance, the S-shaped falls which constitute the rhythm of utterance, deliver the cognitive model at the same time as it (the trajectory) "comes down." Both aspects or modalities arrive at a point of stability at the same point in time.

> Thinking of pitch variation as melody. . .thinking of it as precisely controlled motor action, and then realising that that action is responsible for elaboration of cognitive models gives a new perspective on the S-shaped fall--the creator of rhythm in speech. The point (virtual) is the point (physical). But virtual action is an achievement of a state of play, just as the grasp of a physical movement is the reordering of the world (action) (1977:VII-15).

Mair submits that action and utterance always leave the social reality different from when they start. The reordering of the world may equally well result in a new physical structure, or a new configuration of social relationships. Mair proposes that the world that is reordered by the speech with movement stream is the "text" to that point (1977:VII-14). "Text" is a central concept in Mair's framework. A "text" is a "slice of life": "A section of text is a section of cumulated action along a theme. The theme can be actual or virtual--a topic in conversation, a structure in action" (1977:VII-10). Mair makes reference to a definition of "text" in Bouissac's Circus and Culture which corresponds very closely to his own. Bouissac defines text as "any permanent set of ordered elements (sentences, objects, or actions, or any combination of these) whose co-presence (or collocation) is considered by an encoder and/or decoder as being 'related' in some capacity to one another through the mediation of a logico-semantic system" (1976:126). "Text," then, incorporates all or any of the sensory modalities in the process of communicative interaction. There is also an element of "shared awareness," or "mutual involvement" in Mair's notion of text: "The text is a detailed inter-penetration of two awarenesses organising a shared reality, and the resulting text was their reality in that time" (1977:VII-15).

Mair emphasizes the shared aspect of text in the development of his model. "A conversation is a portion of life text, and during it, the text is shared" (1977:IX-2). He postulates that text is "a trajectory between stable states, and that the transition from percept to plan occurs [for participants] prior to utterance or action" (1977:IX-6). That is, he maintains that text describes a trajectory from stable to unstable states of play, in both virtual (topic) and physical (melody) aspects, and the "plan" of the text is in some sense "in" the brain prior to its manifestation in the speech with movement stream.

Mair says that to view interaction as a system in evolution requires that we look at the trajectory of the system from stable to unstable states of affairs. He borrows from analysis of physical systems, where periods of instability are literally those in which the system is in transition, "where the passage to time can only deliver deliver the system to some sort of stability, even if that stability is mere disintegration" (1977:X-2), and proposes that interaction may be regarded as the survival of texts from stable through unstable to stable states. He argues further, that the "end point" of act or utterance is "marked" by the stability of the system, and that "an incomplete state of affairs constitutes anxiety or challenge" (1977:VII-7). It is in this way, Mair contends, that we can see that shape of melody, the melody of the text, can indeed be interpreted as a sort of action performed on the state of play, and that that action, through its physical shape, organises the other person's awareness. It is from this perspective that Mair argues that production of an S shaped fall in discourse renders stable the physical and cognitive model for all participants (1977:X-12).

> Since speech melody is intrinsic to utterance, and utterance constitutes a spoken text, there must be

> something in their physical form that guides topics to their conclusions, and here it will be argued that in addition to the logico-semantic system (the sign system) of the language, there is a non-arbitrary action performed on the cognitive model in the brains of the participants in interaction by speech melody and it is the shape of the melody which guides topics to their conclusions. . . Speech melody, its pitch and timing, directly wires in the 'state of play' in the awareness of one interactant to that of the other, transforming it not only by what the works so delivered contribute by way of accretion of text, but also by the stability or instability, completeness or incompleteness, of that state of play (1977:X-3,4).

3.3.2 State of Play

"State of play" is a phrase Mair has coined to describe the instant of the ever-advancing now, where consciousness is at in the development of topic during interaction. "It is to be noted that each shared state of play within topic (provided that it really does constitute a shared model in evolution and not misunderstanding) is where consciousness is at that instant" (1978:27). "State of play" is a singular phrase for a complex of notions. "Reality now is always a state of play," Mair says, "even though reality now may be trying to organise reality for another time" (1977:VII-9). At one point, Mair defines the here-and-now state of play as "ego": "The ego is none other than the state of play in the here and now. Each ego is a definition of valid action, an organising centre for and an organised of the social system" (1977:VII-8). Mair argues that "although states of play are private, or in the case of the paranoic, altogether idiosyncratic, the tones of voice and movement still act on them, to deliver the next, and so on" (1977:X-7).

Later, he expands the notion, explicitly including his concept of shared awareness. Mair insists that in interaction the state of play is "always" what the last person said. In this sense, state of play is always seen to be supra-individual.

In one of his more eloquent discussions of state of play, Mair offers a concise, almost cryptic, definition of the term while deftly anticipating objections from more traditional models of interaction analysis:

> It may be argued that the idea of a 'shared state of play', alias possession of the same model in the brain, alias knowing what the other person is thinking is far too mystical and abstract a concept to be of use in linguistic analysis, and even less use in a linguistic analysis which also tangles with physiology. However, I shall argue the reverse, that it is the purpose of the physiological process of speech with movement to achieve a shared state of play between interactants, but that how matters end up--i.e., what that state of play in fact is--often can be sheer misunderstanding and in any case is always where the participants must carry on from (1977:X-7).

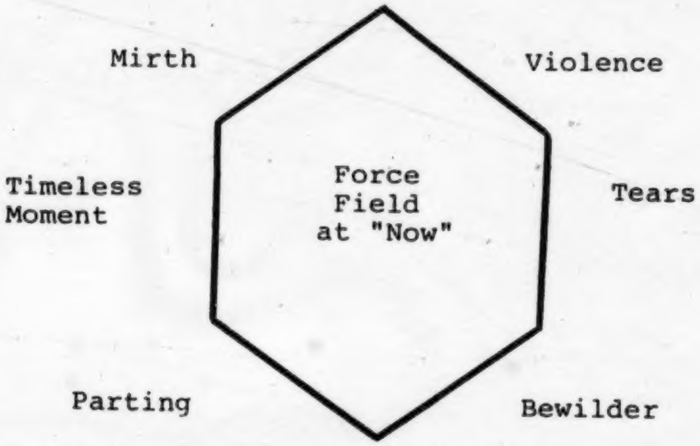
State of play, then, defines a shared model in evolution, the shared awareness of interactants on the ever-advancing edge of time. State of play is the momentary (shared) awareness at the instant of "now" upon which tones of voice and movement act to achieve the next state of play, and so on. Awareness in this scheme is defined as "the path between states of affairs" where "states of affairs" are held to be moments of consciousness composed of the projected, manifested, and perceived plans of participants.

Mair suggests a "decidedly finite scale" for the system which he proposes constitutes awareness at any given time. He refers to the "well known limit to grasp of more than six items in immediate visual awareness (e.g., the face of a dice)" (1977:X-8) and indicates that this limit might prove to be a clue, illuminating the nature, scope, and/or constituent parts of the finite scale system which is the state of play, which is the instant of awareness. Mair posits a "force field" set up by the relationship between elements in immediate consciousness. He suggests that "the experience of the living moment IS the force field of the relationship between elements at the instant of awareness," and further that "the immediately ensuing trajectory IS what this force field resolves into" (1978:29).

The diagrammatic representation which Mair offers is shown in Figure 3.3. Such a notion, he contends, might allow us to reconcile the "rigid matrix of the mutually defining elements of the logico-semantic system," i.e., language, from which texts are constructed, and the "fluidity," the contextual nature, of what it is that in every day life can "come to seem" humourous, boring, sad, and so on.

> Supposing that the instant of awareness is only composed of a maximum of six 'elements'. . .and the relations between them, and the trajectory of the text--the outcome of the predicament, the plan into which it resolves, is the resolution of the force field set up by their relationship in immediate consciousness, then it will be seen that the character of this force field is infinitely variable (1978:28-29).

Figure 3.3--Diagrammatic Representation of Force Field



An axis vertical to the paper = Time = continuing argument (i.e. argument between tropisms)

The "elements" themselves are what Mair elsewhere terms "tropisms." Mair argues that the elements, or tropisms, of the force field at "now" are "long term," part of the synchronic and diachronic cultural scheme into which each individual is born and through which each must make his way. As "creations" of a particular world, they have assigned cultural value. But their effect is not fixed, immutable. Mair argues that in the finite immediate state of play, the instant of awareness at "now," which he suggests is of too short a duration to be caught by introspection, the relations between elements as much as their longer time-scale cultural values may determine the immediately ensuing trajectory (1978:29).

> It is as if the configuration of elements which constitutes each immediate shared state of play has a certain internal tension which resolves into the next state of play, which is then established by

> motor action of either speech or movement (direct action, in both cases, but on a model which is virtual, or physical, or an intermingling of these analytic levels, as is usually the case in human text.) This is the concatenation of relevancies [referred to as the Force Field at 'now'] (1980:23).

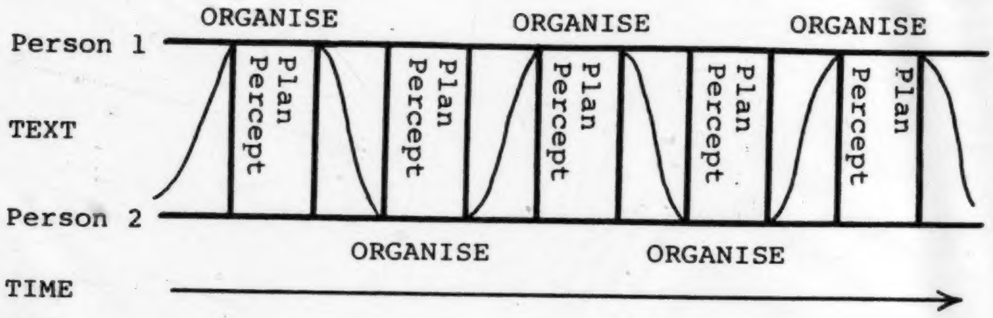
Thus it is the force field of relations between elements in immediate awareness which constitutes the experience of the living moment; it is the force field of relations between elements which resolves into the trajectory by which the next state of play, the next set of relations among force field elements, is achieved.

Finally, with specific regard to the trajectory of text from the "now" of awareness, Mair proposes that "it is the disjunctions in the here and now state of play as compared with the accretion of structured memory which determines the trajectory of the text" (1977:VII-8).

Mair argues that when conditions of mutual involvement obtain, both interactants work at organising the evolution of the supra-individual text. Under such conditions, the brains involved mesh intimately so that the "plan" of one becomes the "percept" of the other. The proposed system may by represented diagrammatically, in stylised form, as in Figure 3.4.

The "melody of the text" is represented in Figure 3.4 by the rising and falling pattern of a sine wave. At each node of the sine wave there is a "percept/organise/plan" sequence which derives alternately from one interactant and then the other. Although the perfect sequencing of turn between interactants portrayed in the diagram may be argued

Figure 3.4--Plan and Percept


ORGANISE
ORGANISE
ORGANISE
Person 1
Plan
Percept
Plan
Percept
Plan
Percept
Plan
Percept
Plan
Percept
TEXT
Person 2
ORGANISE
ORGANISE
TIME

to depict Chapple's form of synchrony, Mair submits that "real life dialogue is seldom like this." He proposes that a "musical taxonomy might best describe the diversity of overlappings, simultaneities, interruptions and ruptures, and so on, each of which is the unique characteristic of a text episode" (1980:17,18). At the same time, Mair insists that the sequence percept/plan manifestation does not occur randomly or without order. He sees text as "an ordered concatenation of plans and percepts" and argues that the text tells a story. There must be and is *relevance* in the ordering of this concatenation" (1980:7).

According to Mair, the stylised presentation here both "improves" and "denatures" the reality turned out by analytic instruments (1980:17). The stylised presentation "improves" the reality in that the up and down melody of the fundamental frequency contours are connected; they perfectly complement one another. The progress of the fundamental through time is thus made more regular, like a sine wave. The stylised presentation "denatures" reality in that there are no phonetic micro-patternings "embroidering" the S-shaped pattern. It is, in Mair's words, a song without words (1980:17).

Mair argues that melody and speech (song and words) are inextricably bound, inseparable components of the natural system of communicative interaction. One melody, he says, is sustained for the duration of any given topic. When topic changes, melody changes, and vice versa. "There is unity of tune over topic" (1977:VII-12). It follows then that the meaning of a melody is part of the state of play which it helps to deliver. Thus, Mair continues, ascription of specific invariable "meaning" to a particular melodic configuration denies the reality. Such ascription demands continual, artificial isolation of melody from the words and momentary circumstances, the state of play, in which it occurred. And such isolation, although appealing in its simplicity, is ultimately folly in its denial of the reality of human interaction.

> Melody is inextricably bound to context.
>
> No one segment of melody can be interpreted separate from the words of which it is a part. That means topic. It is consequently pointless to look outside topic for the meaning of melody. Any such ascription of, for example, an attitude word to a portion of melody is itself retrospective, a time trick, seeking to establish a state of affairs, thus a construct, and with a separate melody of its own (1977:VII-2).

3.3.3 Topic

Several key notions come out of this discussion of the "decidedly contextual nature" of the melody of the text. First is "topic." The notion of topic itself describes the cognitive model in place during interaction. Successive change of topic defines the "cognitive rhythm," argued to be the foundation of the supra-individual system of communicative interaction. Mair's model postulates that in interaction there is an integration of visual and auditory mode processes. The former consist of "plans" of various types; the latter their manifestation in time as motor actions. Mair suggests that "larger scale visual mode plans are thought to be synonomous with 'topic'" (1977:IX-2). Larger scale auditory mode plans are called "task." Both may be described as collaborative and the result of an "accommodation between long range plans of the contributors" (1977:IX-12). In other words, "since a completed topic has one plan, then, if interactive, it is a supra-individual system" (1977:X-8).

> Because topic is also reality for interactants (as organising centres), and because topic is always organising human reality (is constituting it, in fact), then topic elaboration also ends up with a state of affairs (supra-individual) organised in a certain way, and this is the new interpersonal reality (1977:VII-14).

Thus, in Mair's framework, topic is an ascendant descriptor of the process of communicative interaction.

3.3.4 Projection

"The text is a supra-individual system in evolution, and how real-life texts end up must. . .be an accommodation between the projective plans of the participants" (1978:70). Plans are held to be "projective" in that they must have in some sense been "in" the brain prior to their manifestation as coordinated action of the speech with movement stream. "In that a complete utterance is a plan that has been made manifest in time, that plan must have been in the brain in some sense prior to its manifestation in time, and in this sense is projective, that is to say it turns out the immediate future" (1977:XI-5).

Projection in Mair's scheme is argued to be the result (or part) of our ability to perform consummately the time trick, enacting in present linear time a plan formulated at another time. Mair insists that in a true natural science model of interaction we must account for how it is that the future is made into the present, and becomes a shared text in the past (1977:X-10). He refers to the work of another researcher on intonation. Gunter (1974:86) concludes that we always look backwards:

> [W]e cannot project things forward to the response. ... The truth is that we never know enough about the antecedence to predict anything at all about what the next speech will be--or even whether there will be a next speech.

Mair disagrees. Mair argues that Gunter confuses two crucial aspects in his analysis: first, the concepts of projection and prediction; second, the perspectives of analyst and participant. Mair allows that Gunter is correct in concluding that a statistical study of probabilities of a certain utterance following other utterances would not be useful in an attempt to understand the principles of text regulation, but, he proposes that "in spontaneous interaction we do in fact project things forward, project topics at their destinations, just as the hands project the plan in the brain, beamed through the eye, onto the physical medium" (1977:X-11).

Mair is not unsympathetic to Gunter's predicament. He has also found "retrospective predictability" to be perhaps the "most disconcerting insight gained from prolonged study of real-time texts" (1980:24). He escapes the impasse of that insight, however, with the concept of projection.‘ Mair postulates a mechanism by which topoics are projected at their conclusions--the fundamental frequency contours and movements turned about by single central patterning programs in the brain--and suggests that it is the action of the mechanism upon the state of play in the continually advancing "now" which projects topics at their conclusion (1978:69). He argues that interactants have an "awareness" of the mechanism such that they can and do predict when in time the physical trajectory of utterance will "come down" on the S-shaped fall of the melody, rendering the physical

---

‘The concept of projection allows Mair to escape from the impasse of retrospective predictability, but, as he notes, it "opens up the epistemologically more serious problem that the theory now has with the concept of 'dimension'" (1980:25).

model stable and delivering the semantic point of the utterance.

3.3.5 The Linearity of Awareness

This notion is implicit in much of the previous discussion, particularly the elucidation of Mair's description of reality as culturally constructed, a "mask and support for a certain kind of society," something in which human beings are thought to be totally immersed by virtue of their ability to perform the time trick consummately. It is the relationship between the linearity of time, the linearity of awareness, and the ability of an argument to reverse or backtrack which Mair seeks to explain with this portion of his model.

Once again he begins with a notion "so obvious as to be invisible": time is unidirectional and linear. Each moment in time, then, is different, one of a kind, never to be repeated or replicated. Mair proposes that the time trick seems to "halt" the unidirectionality of time and that it is successful "to the extent which the plan works, and time is suspended, or circular" (1977:VII-13). The notion of linearity of awareness comes from Mair's consideration of two modes of communication in which awareness must be linear for that portion of the text: touch and music. Touch, he says, is informed by the shape made by the other. "A caress which does not conform in shape to the actual presented contour loses what one is tempted to call magic, but must now be called simultaneous co-patterning of person other than self" (1977:VII-21). Music is also a mode of experience where awareness must be "in" it for it to be experienced. "Music is interweaving trajectories, and thus must be experienced on the run. To stop listening long enough to reflect. . .would be to miss a bit" (1977:VII-22). With music, as with touch, awareness must be linear during that part of the text "or one is no longer there."

From this realisation of the "nature" of touch and music, Mair posits linearity in both acoustic and visual modes. First he argues the acoustic mode must be linear:

> The auditory mode consists of time forms, and these are necessarily linear. ... There is but one larynx per person, and this can only modulate one fundamental frequency. ... One cannot be many. If this is so because it is merely one of the logico-aesthetic tendencies of the human mind, it is no less true for that (1977:VII-3).

Then he documents a "linearity" of the visual mode.

> The visual mode consists of organised time forms, which hold stability over time and can be considered as manifested plans. Either they move, or the eye moves over them, but in following, or in fixating, the result is to temporally immobilise an image on the retina long enough to permit recognition (1977:VII-3).

Finally, he argues that both acoustic and visual modes are typified by a process in time, "scanning," which is segmented by plans, themselves timeless, constructions in the brain.

> In both visual and auditory modes we have a process in time, scanning, which is segmented by plans which are timeless. Whether that which is timeless is built into the structure of scanned objects, or is the result of structuring objects (actions), in which case the plan must have in some sense been in

> the brain in the first place, or is there in memory as an accretion of the time form of an utterance, or as the plan by which an utterance was structured--whatever manifestation requires plan and scanning (1977:VII-4).

In sum, time is unidirectional and linear; action and utterance segment time into events, chunking the continuum of reality; and, "at any one time awareness is only somewhere" (1977:VII-1). Awareness itself must be linear.

Mair proposes an identity between the abstract theoretical concept of the linear trajectory of systems between stable states and the actual mechanics of consciousness. "Awareness itself is the linear trajectory between stable states of achieved or perceived plans" (1977:IX-11). He argues that "at any one time, a person is his or her awareness, the linearity of which it is difficult to catch sight" precisely because it is linear (1977:X-8). To catch sight of it is to step outside it and glimpse the system "on the run" as it happens. But the constraints of the system, the constraints of time, make that (at the very least) problematic. Reconstruction of a past state of play requires replication of that moment; and that, to date, is impossible. Mair argues that "retrospective reconstructions of states of affairs, whether performed for psychological theories, or as part of a later state of play, will only select those bits which can 'make something' of the later state of play" (1977:VII-13).

Each time slice is unique and the text must carry on from where it has got to in time. Thus we cannot step "outside" to observe, evaluate, or reconstruct past states of play with the objective impunity which we have often assumed. We are constrained by the force field of elements which constitute the moment of awareness at "now"; we are caught in the linearity of time. Mair argues that the attempt to attribute verbal descriptive categories to past states of play is itself "retrospective, a later portion of text with its own melody" (1977:VII-19). Such attribution must, in the process of reconstruction, create a later state of play. Thus, to say that an argument can backtrack, re-turning upon itself to revisit former topics is in fact to say that a current state of play, "now," that is concerned with specific elements of a former state of play has been created and is in the process of unfolding.[7]

> To find words adequate to portion of of past experience is to try to re-create that state of play of that past text is always a reconstruction, and it is suggested, is more nearly approached with a poetic use of language than by any restricted assembly of operationalised concepts (1978:71).

It is from this perspective that Mair proposes the "context-bound relativity of effect" held to typify melody. If each time slice is indeed unique, and the text must carry on from where it has got to, then it is pointless to search for generalized, inviolable rules governing the effect of melody in interaction. Melody acts directly on the brains of interactants, transforming their immediate awarenesses. Interactants themselves are immersed in the evolving model. Receipt of a given down- or up-going melody must then depend on the interactional situation of the moment, or context, for its effect. There are and can be no immutable laws governing receipt or projection of specific melodic configurations.

---

[7] In his review of Bouissac's Circus and Culture, Mair summarizes the difficulties imposed by this particular aspect of time and awareness on interaction analysis, presentation of theory, and theorising in general. His remarks apply equally well to this presentation of his theory: "It is the fate of texts to be linear, even texts of texts. . .The peculiarity of the auditory mode, and perhaps attention itself, is a great handicap in a 'proto-theory' such as this where each element depends on the others. A linear text such as this could weave many paths through the issues we have at present, and the particular order in which they are here presented is only one (1977:VIII-3)."

3.3.6 Summary of Rules Governing Speech Melody

Despite the "context-bound relativity of effect" that characterises intonation, Mair argues that it is possible to detail alternatives open to the other upon receipt of a down-going or up-going melody. He posits three rules which govern the interplay of speech melody, state of play and projective plan (1977:IX-8; 1978:25):

> 1. Within topic, a down-going melody delivers a state of play that is intrinsically stable. A down-going melody is a further linear development of the topic. If this is in accord with the projective plan of the other, then this is harmonious topic conclusion.
>
> 2. If it is not in such accord, then an utterance with a down-going melody that follows it will be either new information--an extension of topic, or a confrontation. If topic is maintained this escalates to violence.
>
> 3. An up-going melody delivers a state of play that is intrinsically unstable. [An up-going melody] following a down-going. . .backtracks the text, puts the state of play back to where it was before the immediately preceding down-going. Within topic, it

> constitutes something that must be worked through before the text can move on.
>
> (An up-going will be followed by a down-going, if that topic can proceed.)

It is in this manner that shapes of melodies are held to act as topics, guiding them to their conclusions, the text taking on the form of a trajectory moving from stable to unstable states of play. It is from this perspective that each interactant may be seen as part, a node, of the supra-individual system in evolution.

3.3.7 Immersion

Mair argues that, under conditions of rapport, interactants are literally immersed in the evolving model which is held to constitute their awareness at any given moment in the evolution of the text. "The fact that little movements and sounds trigger shared realisations implies that everything up to those shared realisations was the same for the participants" (1977:VII-13). The physical character of sound seems to support Mair here:

> All the participants in a conversation can be considered immersed in the speech rhythm of the speaker, since, as Carpenter and McCluhan put it: 'the essential feature of sound. . .is not its location, but that it be, that it fills space' (1977:VIII-22).

Even a written text, Mair suggests, imposes the "cognitive and implicit physical rhythm" of the author on the reader when it is read.*

---

*The effect of a written text on its readers when read has already been discussed in Chapter 2. To reiterate, Zlatoustva (1975), among others, has found that speakers, Mair proposes that in conversation, as in the culturally constructed world of everyday reality, there is a paradoxical relationship between immersion and obviousness, or clarity. ("The other side to Immersion is Clarity" [1977:VIII-22].) "Clarity" refers to that which is so deeply assumed that it is obvious: "For something to be obvious, it must be deeply assumed. And the deeper the assumption, the less visible does it become that reality could have been otherwise" (1977:VIII-22). Mair considers "immersion" to be the thing that allows for the progressive condensation of meaning typical of interactions where the conditions of "rapport" obtain. The shared history of interactants permits progressive condensation of meanings. "Less and less can mean more and more" precisely because interactants share references, implicit knowledge of past shared states of play.

Mair describes interactants, immersed in the evolving model (at both the cultural and conversational levels) as "caught" in a "bind" of reciprocal control. He links his concept of "immersion" with that of "projection." The state of play, the model in the brain, the instant of awareness, is shared but the projective plans from it (the moment of shared awareness) may differ. Interactants must, then,

---

[*] (cont'd)when presented with a text from which all graphic regularity and indicators of rhythm had been removed, could readily determine the metrical quality of the passage from the sequence of regular rhythmic structures in the passage and would adapt their speech pattern to it. A written text may thus be argued to impose its "cognitive and implicit physical rhythm" on the reader, when read.

accommodate one another's projective plans; they must "compromise" or the system, in this case their entire momentary awareness, may disintegrate. Thus interactants "control" one another.

3.3.8 Summary of Principles

From these notions Mair (1978:22-23) derives "three basic principles" of text regulation:

> 1. Speech structures reality.
>
> 2. Speech and movement are turned out by the same patterning program. Speech and movement are so closely coordinated that it is more economical to assume that they are turned out by the same patterning program. The data show numerous sound and movement clusters, bursts of activity including vocalisation, head movement, facial movement, even manual gesture. These bursts of activity are the property of the dyad, and the temporal integration among the interactants includes such phenomena as coming in together after silence, smooth sequencing of repartee, actually moving synchronously as in the onset of enraptness.
> . . .
> But just within one person, this co-patterning is more than just the precision of temporal integration. The movements, in context, actually seem to be a partial manifestation of the ideas of the segment of speech with movement of which they are a part.
>
> 3. Immersion in the evolving model. Immersion in a shared history of the conversation permits progressive condensation of meanings--less and less can mean more and more, but it depends upon immersion in topic, i.e., shared history prior to the condensed communicative event. Immersion in a shared visual here and now also permits communicative events to be condensed.

3.4 Rapport and Control

Closely related to the concept of immersion is the notion of rapport. Rapport is a fundamental part of Mair's framework. Most of Mair's model has been derived from intensive analysis of audio and video tapes of interactants "in rapport." Mair's discussion of the model has been with data that are collected when the interactants are in a state of rapport. Of necessity then, a good deal of the discussion above invokes the notion of rapport, assuming an implicit, shared "common sense" meaning of the term. Mair defines rapport as "immersion in the here and now." He argues that rapport is "a mode of experience into which individual conscious is lost" (1977:IX-7). This is similar to the broader category of experience Mair calls "entrancement." The experience of entrancement, according to Mair, seems not to be limited to the physical process of speech, but is part of any activity in which one is "enrapt." To be entranced or enrapt is to take the timing of one's own cognitive processes from the rhythms impinging on the senses (Mair 1978). It is the characteristic of "neuroplasticity" which is discussed by Bouissac and others in this context which allows for the the take-over of brains by external rhythmic sources. Thus, to be enrapt, to experience rapport in interaction, is to be totally immersed in the here and now of interaction, abandoning "self" to the extent that one is taking the timing of one's own cognitive processes from the rhythms of the speech with movement stream which all interactants work to establish and to organise. To be in rapport is to move "in time" (behaviorally and cognitively) with one's interactive partner(s). "If one is 'caught' by one's partner (i.e. enrapt), then the trajectory of his or her utterance is also predicting the point, so that they are moving together, in time" (1977:VIII-25).

Mair explicitly connects rapport to two essential aspects of his model. The degree of synchrony in interaction is purported to be contingent upon the degree of rapport between interactants. Mair argues that evidence of synchrony, and rapport, is found in all types of interaction, but that some interactions tend to have "more," in terms of both frequency and intensity, than do others. Both rapport and synchrony seem to be somethings that interactants can "slip into and out of, as it were" (1977:IV-14). The "cognitive rhythm" of natural conversation--which Mair proposes as the source of synchronous movement and speech, and the smooth sequencing of events in interaction; and from which Mair originates his notion of a single central patterning program in the brain--is also closely allied to rapport.

> The timing of cognitive change points, both of perception and transformation, has a highly variable relationship to the physical process of utterance, except under conditions of rapport, where it will be argued that awareness is 'in' the text such that the timing of the physical process has a systemic relationship to the change points in cognition (1977:IX-6).

It is the linking of the notion of projection to the notion of melody as hard-wiring, and the combining of these notions with that of rapport that is the kernel of Mair's theory of text control.

Melody is shared, supra-individual. The path of the melody through time describes a trajectory. The falling S shapes deliver the cognitive model, rendering it stable at the same point in time as they physically "come down." Melody "hard-wires" the consciousness of interactants, controlling the direction of discourse. Under conditions of rapport, the immediate awarenesses of interactants are "conjoined. . . .[so that] we become as others are, and seek to make others become as we are in time" (1977:X-16). Thus, he who controls melody controls text, controls topic, controls interaction. Once gained, control of melody is argued to "ensure the predominance of its progenitor in the way the text develops" (1977:IX-7). In other words, within the model is the assumption that interaction is a collaborative effort, an accommodation between the projective plans of participants, but the one who ends up with interaction following his design most closely is the one who accesses and controls the melody of the text. In more traditional terms, he who controls the melody may be characterized as the one with "charm" or "charisma." "It is a concept more Machiavellian than emotion--text control is the outcome--and at the same time assuming a connectedness which the notions of emotion and display rules lack" (1977:X-16). Interaction then may be viewed as a sort of melodic competition for control. Mair summarizes his theory of melodic control of text, (re)placing it into the "real world" of human interaction:

> In interaction, there _is_ always continuity in the melody across participants, and a winning argument, unless asymmetric status relationships have skewed the bias, _will_ always recruit a melody asymmetric to and which completes that of the vanquished opponent (1980:24).

A corollary of Mair's theory of text control is the notion of "cognitive survival" and/or "risk" in interaction. He argues that each utterance puts at risk the future of the interactant as organiser of text. Each utterance is a gamble of sorts for interactants; each utterance risks not only rupturing the shared model which constitutes their momentary awareness, but destruction of the credibility of the utterer as text organiser. Interactants "battle" in a sort of melodic competition for control of melody, control of text. The other must carry on from the first. What happens next is determined by whether or not the state of play delivered by the first is in accord with the projective plan of the other within that topic. At each turn, then, interactants risk their own cognitive survival as credible contributors to the evolution of the text. Mair calls it an existential predicament: "Each state of play within topic is. . .an existential predicament, and the ensuing melody delivers the next, and so on. It is a shared predicament, although at any one time the outcome projected by each may markedly differ" (1977:IX-7). Interactants gamble all with each utterance. It is the nature of the system.

Mair also interprets pitch range and excursion in terms of text control. He refuses to partake of the "interpretive variable approach" commonly associated with analysis of pitch in interaction. In what Mair calls the "popular understanding" a flat or monotone voice is considered indicative of uninvolvement, depression, boredom, and so on. Voices with more fluctuation and higher pitch, on the other hand, are thought typical of excitement, interest, involvement. Mair proposes that pitch excursion--"going over the top of the melody"--more properly describes a bid for control of topic.

> The suggestion is that. . .the stability or instability of a cognitive state of play--the obstacle that one must overcome in shifting it for example--does not only depend on the immediate contextual and longer term cultural values of the elements of its constitution, but also on the range in pitch with which it is delivered (1980:24).

He suggests further that to "rescue the point" of a topic that

> . . . is going against one's own projective plan--to deliver it back to one's own projected outcome, needs progressively higher arousal, [progressively higher and stronger pitch excursion] the more it has already diverged from that outcome (1977:X-22,23).

If the topic continues to go the way that neither participant had projected, Mair documents a continuously escalating melody of confrontation and eventual eruption into violence that is consistent with the function attributed to pitch range and pitch excursion in his model.

Mair says that the relinquishing of control in interaction is accompanied by "vulnerability," whether that is the outcome of melodic battle for control or an abdication from text control by one participant or the other--or even the more subtle, and perhaps gentle, "wooing" of control from one by the other. He describes the process of giving over control to another as an aspect of "letting go" in interaction, letting one's own awareness take its timings from the other's rhythms of speech and movement.

> As with touch, the sharing with another of one's own motor and sensory processes, it [the giving over of control] is accompanied by vulnerability. It is, after all, a portion of life-text that is being generated, and to give over control to another is to submit to transformation with uncertain outcome (1977:IX-12).

3.5 Tropisms

Mair considers a finite number of ways that topics may end up in interaction. He applies the term "tropism" to these "final common pathways"[1] to emphasize that they are "behaviour patterns" that are extremely difficult to "abort" once the shared text has embarked upon them (1978:32). He argues that tropisms are patterns. They are "sequences like laughing and crying which, as motor patterns, appear to have their own direction and time course" (1977:X-13). Tropisms, then, are a bit like "fixed action patterns" in ethological descriptions. Mair says that "to 'pull' a tropism in the course of a text is to determine very powerfully how that text turns out" (1978:33). He outlines his tropisms thus.

> Harmony.--the complete model, the resolved problem, the solved equation

---

[1] The term is taken as metaphor from neurophysiology.

> Violence.--(or the continuation of the text by other means)
> Parting or interruption.--the argument left unresolved
> Tears.--complete fragmentation of the immediate state of play. Unlike the cube flip, in which both interpretations of the immediate state of play are co-present, or the logic of the developing argument, in which states of play successively become the case, with the tears tropism there is no ensuing text generation because there is no immediate state of play. Others must move into that consciousness to deliver it back to viability. Since they will spend time doing this, the tears tropism is a very powerful weapon of text control.
> Timeless Moment.--the realisation that the other person has realised that one has realised that. . . etc., (to an infinite regress). The timeless moment has perhaps a special status: whereas the moment of coming into existence of a shared state of play in dialogue usually has content, the realisation of togetherness in time itself has no content. The topic-free togetherness that is the timeless moment cannot endure for long (nothing is fundamentally happening during it) but from that point of silence the intersecting text of interaction can radically change its course.
> Bewilder.--the state of play becomes an impossible model like Escher art-work. . .[and] this halts the text, and puts the interactants 'back in boxes.' They must start again (1978:33).

As noted above, Mair's tropisms are the "elements" that compose the force field at "now," which is the immediate awareness, the shared state of play of the interactants. The tropisms are represented diagrammatically in Figure 3.3, in the explanation of Mair's notion of the "force field."

Mair says that the "configuration" of elements constituting each immediate shared state of play has an "internal tension" that resolves into the next state of play, which is itself established through the direct actions of sound and movement on the immediate awarenesses of those involved. However, if a state of play violates the limits of instability or over-stability, then, Mair argues (1980:23), a tropism ensues and there is no more "concatenation of relevancies" to produce the "internal tension" characteristic of normal, typical immediate states of play. Mair suggests that there is a specific, culturally- and contextually-defined range of possible "commitment and energy" within which normal interaction can proceed. To move beyond that range is to embark upon a tropism.

The notion of tropisms, or final common pathways of human interaction, is a difficult portion of Mair's model. Perhaps it is better to let Mair explain in his own words:

> One can say that normal interaction can only proceed while the individual is operating within a defined range of potential commitment and energy, the actual position in this range for any immediate state of play being cultural and contextual; and that normal interaction is the experience of an individual in that range as manipulated in and by the text. The text moves the individual within the physiological range. Where the text itself moves the individual outside that range, we get Tropisms. But for an individual whose system is already outside that range, the text is one long tropism. . .(1980:26,27).

3.6 Summary

The theory has to do with the relationship of cognitive and physical process in interaction. Spontaneous conversation results in what Mair terms a "shared text with a logical structure." Mair is concerned with definition of the "mechanism" by which interaction proceeds. He aims to account for how it is that the plan of the topic is worked out in the present and becomes a shared text in the past.

Specifically, he intends to delineate the mechanism by which interactants "project" topics at their conclusion; that is, how they manifest in time plans that have in some sense been "in" the brain prior to their manifestation as the speech with movement stream.

From his intensive analysis of audio- and video-taped conversations, he proposes that fundamental frequency contours (speech) and motor action (movement) are patterned by a single central patterning program in the brain; that these two elements or modalities act to guide topics to their conclusions; and that the central patterning program manifests itself in the various forms of what Mair terms intra-personal synchrony. The system, the mechanism, is represented diagrammatically in Figure 3.4.

The rising and falling S shapes which link the percept/organise/plan sequences are stylised representations of the speech melody, the creator of rhythm in speech. Speech melody is supra-individual, demonstrably shared. Mair argues that speech melodies are not "signs of" external variables, but are, rather, direct actions on the cognitive models of interactants at the time of their (the melodies') occurrence. Similarly, little movements accompanying speech are not merely "actions" or "illustrators" accompanying the semantic import of the "state of play" they help to deliver. Head and eye movement is held to be particularly noteworthy. Melody and movement "hard-wire" the consciousness of interactants, performing "non-arbitrary actions" on cognitive models in the brain, controlling the direction of discourse. The S-shaped falls of utterance, the trajectory of utterance, deliver the cognitive model at the exact same time as they "come down" in time. The acoustic (physical) and cognitive (virtual) model are rendered stable coinstantaneously. "The point (virtual) is the point (physical)," particularly under conditions of "rapport."

The physical and physiological constraints of the system demand that melody be characterised by a "context-bound relativity of effect." Melody and topic are welded together. No one melody can be interpreted separate from the words and movements of which it is a part. Mair documents a unity of tune over topic: when one is changed, the other changes. Melody is inextricably bound to context. At the same time Mair posits three rules governing the action of melody in and on natural conversation, detailing the mechanism by which topics are guided to their conclusion:

1. Within topic, a down-going melody delivers a state of play that is intrinsically stable. If in accord with the projective plan of the other, this constitutes harmonious topic conclusion.
2. If not, then a following down-going melody will either be new information, an extension of topic, or confrontation.
3. An up-going melody delivers a state of play that is intrinsically unstable. Within topic, it backtracks the

> text; it is something that must be worked through before topic can proceed. If topic can proceed, an up-going will be followed by a down-going melody.

In this manner the text takes on the form of a trajectory from stable to unstable states of play. Mair proposes that the text to date is always the trajectory by which it got to where it was. The time trick enables planning, and construction of the physical world as "mask and support" for the societal model, as well as the (subsequent) immersion of our species in its various constructed worlds.

The model postulates that interactants in a state of rapport share awareness. This is perhaps the key notion. Mair contends that participants in conversation must be considered as parts, nodes of a supra-individual system in evolution which they themselves, by their very nature and by the nature of their biological structure and organisation, create, define and maintain. Under conditions of rapport interactants are thought to be immersed in the evolving model which itself is argued to constitute their momentary awareness. Individuals are organising centres for the state of play, the evolving system, which itself consists of those individuals as nodes of a supra-individual organisation. Interaction is, then, systemic; it is patterned and rule-governed.

Both (or all) interactants work at organising the supra-individual topic. The shared melody serves to project topics at their conclusions, and participants, immersed in the evolving model, are able to predict when in time it (melody, and consequently, the entire model) will "come down." Furthermore, in the process of unfolding the supra-individual model interactants evidence synchronies of all sorts. There are intricate coordinations of timings and movements, between interactants, and within and between modalities. Speech and movement are shared between interactants. This sharing is termed inter-personal synchrony.

In sum, the model proposes that interaction is a supra-individual system in evolution, producing a shared text with a logical structure. Of necessity the text represents an accommodation between the projective plans of the participants. The model is shared, but projections from the shared momentary awareness may differ. Interactants are thus thought to be caught in a bind of reciprocal control. Accommodation is necessary for survival of the system. What is more, the concept of individuals as organising centres joined up by sensory modalities and immersed in the evolving model upon which melody performs direct, non-arbitrary actions, suggests that he who controls melody controls text. He or she is thought to be the guiding force in the way reality is turned out. Faces, movement, speech, intonation contours are held to be the "forces" which actually construct this reality, the reality of human interaction. Mair says that they act on the evolving model, and he proposes that words to describe those "forces" are retrospective--"emotion words"--part of a later state of play, description, with a melody of its own. In sum, the theory would have it that "in conditions of rapport, there is a shared awareness and that tones of voice and movement act on it to control outcome" (1978:73).

Chapter 4

Critique of Mair's Model

4.1 Introduction

Mair is still developing his model. The appeal in the theory developed thus far has to do with criteria of descriptive, analytic, and observational adequacies. The use of language takes place in a social context; it is humans in groups who speak and understand. Other language-focus models, in particular context-free grammatical models, have failed to account adequately for contextual variability in speech and other communicative performance. Mair's Steps Towards Principles of Text Regulation proposes a model that allows us to at least begin to seriously consider the process of human interaction ~~as~~ it happens, in the context which it organises and by which it is organised. What is more, his model addresses some of the more theoretically obscure (obscure in that they have not been amenable to description, analysis or explanation in the past), yet intuitively obvious and indispensable aspects of ordinary talk. In Kuhnian terms, Mair proposes a paradigm shift, the ramifications of which resonate far and beyond any specialised study of language.

Mair's model is not without its limitations, however. The previous chapter has presented a "bus stop" version of the theory, documenting the growth and change from an initial concern with faces to the inclusion of speech and gesture, and the broader, all-inclusive theory.

In this chapter there is a more critical discussion of Mair's model, exploring specifically those aspects of the model which are relevant to the present study's concern with the phenomena of synchrony in human interaction. A substantial portion of this segment of the text is devoted to the introduction of a theory of observation of natural systems recently proposed by H. Maturana and F. Varela (1975, 1976), which has major implications for Mair's model and the study of interaction in general.

Mair argues for the interrelationship of all aspects of the speech with movement stream. He is also cognizant of the necessity to consider all elements of a natural system in the contexts in which they occur and of which they are originally a part. His model is, consequently, "transmodal" or "transmedia" in nature. "When the eyes go precisely with these minute voice movement clusters. . .one begins to suspect that any division of the speech with movement stream is arbitrary (or conventional)" (1978:33). At the same time, he insists that interlocking aspects of human interaction must be demarcated--for analysis and explication--or "all we will have is soup" (1978). Furthermore, Mair is well aware of the difficulties of validating chosen units. Specification of criteria for the choice of units forms an essential part of his framework.

Mair's own fundamental units, the five "artificially extracted parameters" attempt to satisfy his own definition of natural science criteria. They are observable, largely quantifiable, and, it may be claimed, biologically-based variables. They are what Mair has called a "best alternative." These parameters allow us to talk with a large measure of descriptive, analytic, and observational adequacy about the totality of the process of human communicative interaction.

They work, apparently. From Mair's artificially extracted parameters it is possible to formulate description, analysis, and explanation which seem to be very nearly adequate to the perceived and intuited reality of the complex process of interaction. They are, nevertheless, admittedly the result of more or less arbitrary decisions and represent an arbitrary segmentation of the data. They may (or may not) approximate parameters actually employed by participants.

Mair's criticism of past interaction research holds. That he finds his units useful may reflect his own interpretive bias as much as the descriptive, observational, or analytic value of the units; that we find his units useful might only reflect the precision of culturally coordinated decision making. Mair has not found a clear way out of the tangle of foggy notions that surround investigation of the process of communicative interaction. Specification of criteria for definition of units and for boundary demarcation remain difficult problems.

4.2 Descriptive Adequacy

Mair posits a "mechanism" of interaction. He maintains that although states of play are "private," tones of voice and movement act on them to "deliver the text," guiding topics to their conclusions. Melody and movement comprise the mechanism. Several aspects of interaction, several observations on interaction may be accounted for with this fundamental concept.

1. First, interpreting the effects of speech melody as action performed on a visual mode "now" greatly simplifies accounting for the way topics are guided to their conclusions (1977:IX), and in so doing, satisfies a basic dictum of Mair's natural science model, Occam's Razor. If speech melody and movement do in fact act directly in immediate awareness, continually transforming it, there is no further need to invoke the host of other parameters which inhabit the interaction literature. Identification, description, and analysis of the "mechanism" become paramount.
2. Thinking of melody as action obviates the need to operationalise the "feeling words" available in a language as part of a theory of text regulation in communicative interaction. It avoids the problems of para-linguistic researchers who try to operationalise descriptive words for experience. As noted above, it is precisely as researchers attempt to "operationalise" their concepts for feelings, attitudes, and emotions,

that they "lose the very adequacy of their concepts to lived experience" (1978:29).

3. Mair's proposed mechanism obviates the segmental versus suprasegmental dichotomy in linguistic analysis of communicative interaction.
4. Thinking of melody as direct action on the brain provides a specifiable function for short-term memory, while situating interaction firmly in the fleeting context of "now."

   > If we see little movements and pitch as direct, literally building the text as two people might assemble a building, then we have a function for out separate short-term memory system, and a way out of having to look for either standard forms or manifested categories for pitch change or movement. Descriptive building trade terms like beginning, ending, almost rounding off could describe the contours, but interpretation of what they meant' would have to look to the whole conversation and its context for full resonance (1977:VII-17).

5. Thinking of state of play in awareness as a visual mode process also connects it to the "now" of its occurrence, emphasizing that awareness (consciousness) is always "now," "a physical process like any other in this respect" (1977:IX-5). To describe state of play as a visual mode process requires, in addition, that the change points between stable states of achieved or perceived plans (which plans, under conditions of rapport, are held to constitute awareness) are thought of as abrupt, like the "flip" of a Necker cube.
6. Thinking of melody as action in linear time accounts for the "context-bound relativity of effect" by which

intonation is characterised, while obviating the need to present formal, TG type rules for interaction. As Urion (1978) argues, in order for a formal rule to be effective, a formal context for application of the rule must be specified. "That is equivalent to a definition of the 'structural description' of transformational grammar, which specifies structural context for application of a specific rule" (1978:198). Furthermore, as Urion points up, while the specification of context for rules of utterance may be an attainable goal, specification of context for rules of interaction becomes an unmanageably cumbersome task. There are too many modalities requiring context-rule specification at the same time. A rule which would allow and account for violation of one set of rules at the expense of another is required. Urion (1978:198) concludes that "instead of addressing the coherence of sets of rules as the descriptive problem, it seems more nearly adequate to accept structural regularities in both ... [speech and movement] as both given and contextually manipulable." Or, as Mair argues, we must look at patterns of interaction as they are established by communicative participants; we must investigate the system of interaction, keeping in mind that the "observation that behaviour is not random does not immediately prescribe the definition of pattern as 'rule'" (Urion 1978:198).

7. Moreover, the context-bound relativity of effect of

melody "may account for the infinitely variable retrospective reconstructions of the experience of words (discourse) that people attempt" (Mair 1977:IX-8).

4.2.1 The GT Model

Mair derives his unique notion of Gestalt Transform (GT) from such notions. A gestalt is composed of an amalgam of constituent components. A facial gestalt, for example, is composed of a "blend" of muscular contractions. The GT model predicts that minute changes in gestalt constituents will not necessarily change perception of the gestalt--they are by definition imperceptible shifts--until a "sufficient" number of constituent changes pushes the perception of the gestalt to a breakpoint, and a new configuration is perceived. Mair proposes use of the gestalt model at both higher and lower levels of analysis. The higher level the gestalt, the greater the diversity or potential diversity of its component parts. Lower level gestalts such as "face" (as in "to make face") are thought to be composed of a variety of components and component types, and these may include such things as functional information.

The boundaries of a gestalt are by definition neither static nor precise; it follows then that the boundaries of higher level gestalts, which are themselves argued to be composed of lower level gestalts, may be susceptible to influence and change. Furthermore, a change in a higher level gestalt must affect analysis of its constituent components, but the converse is not necessarily true. A change in a lower level gestalt, considered as component of the higher level gestalt, may not affect perception, and should therefore not affect analysis of the lower level gestalt. The GT model thus demands that interaction analysis take account of context.

The GT model also has some important implications for sociolinguistics (after Bullard 1980). Sociolinguists propose that minute, imperceptible changes in a sound which occur in the process of verbal interaction are the main vehicle for sound, and thus language, change. Labov (1979) has begun to focus attention on the interactive setting as both the locus of and vehicle for change. He argues that the cumulative effect of independent, successive interactions eventually produces a shift. Language change may thus be construed as a GT over an extended period of time. In this manner, the GT model may be held to account for synchronic and diachronic processes.

4.2.2 Immersion

The notion of "immersion" in a "shared awareness" is a key notion from the theory. In conjunction with the concept of speech melody and movement as direct action on immediate awareness, "immersion" may be argued to account for significant aspects of the process of interaction. To consider interaction as typified by participants immersed in a shared model in evolution allows theoretical room for the often experienced feeling of "strangeness" in a discourse with someone who is "making of the physical events of the text an entirely different text to one's own" (1977:X-22). Mair argues that we do not need the long-standing linguistic dichotomy of affect versus message to account for the experience of strangeness. The complex of notions involved in "immersion" and "rapport" seem to more than suffice. Mair reduces the grammar versus attitude dichotomy to redundancy and notes that his interpretation of the process of human interaction receives support from the "non-trivial" observation that

> . . .the strangeness [of discourse with someone who is making of the physical events of the text an entirely different reality to one's own] only disappears if such a person slips back into the same here and now as the other with whom he converses (1978:34).

In addition, ordinary every day control of discourse, "including the control of information in dialogue (topic control)" (1977:X-7), and the common experience of losing one's self into an interaction are explainable as part of the "effect" of Mair's proposed "mechanism" of interaction (i.e. speech and movement) as it acts upon interactants immersed in the evolving model.

4.2.3 Projection

Tones of voice and movement are the means by which topics are projected at their conclusions: they are the mechanism. Their action upon immediate awareness in the state of play at "now" accounts for how it is that we project topics at their conclusions in everyday discourse. Mair argues that not only do interactants have this ability to project topics at their conclusions, but they can and do predict when in time the physical trajectory of utterance will "come down" on the S-shaped falls by which it is rhythmically structured. Interactants do in fact predict the points of stability of the system, i.e. the points where the cognitive and physical models "come down" together. And this speculation, itself dependent on the concept of immersion in a shared reality, might at last provide an observationally adequate explanation for the detailed choreography of synchronous and sequential speech and movement in spontaneous interaction.

> The close study of conversation suggests that interactees are not only making a shared melody, but that the physical trajectory of utterances (which constitutes their rhythm), is predicted by the other interactee as to the place in time where they will 'come down'. ... [This postulate] might be an adequate explanation of the synchronous movement and smooth sequencing of communicative events among contributors to a 'natural' conversation. (1977:VII-24).

4.2.4 Charm and Charisma

Mair suggests that notions such as "charm" and "charisma," previously unamenable to scientific investigation, may be accounted for in a framework that examines interaction as a supra-individual system in evolution, where the action of melody upon state of play performs direct transformations on immediate consciousnesses. He describes charm as a skill, controlled manipulation of the melody of the text. "Despite the fact that sweet talk can only act on "now," certain professionals such as doctors, hypnotists, and ritual agents of all sorts can develop it as a skill, but this is the same as developing the personal quality of charm" (1977:IX-12). Mair submits that his theory of text control may explain how it is that one can take over, or be taken over so completely during sweet talk (under conditions of rapport, when entranced). What is more, he proposes that the model may be extended to account for the phenomenon of hypnotism. "From this perspective, telling someone to do something and a result of them in fact doing it is seen as different only in degree from the phenomenon of posthypnotic suggestion" (1977:X-23).

4.2.5 Ostensible Versus Oblique Topic

Finally, the notion of melody as action might allow us to account for the frequently experienced discrepancy between what is actually said and what is really meant, between the literal meaning of words and "what we know that the other person will know that we meant." As Mair speculates, "it (the 'mechanism') might allow us to think seriously about the proposition that for many of our most important shared realisations between people, nothing need be said at all." (1978:18). To be in rapport is to be entranced, enrapt; it is to take the timing of one's own cognitive processes from rhythm's impinging on the senses;

it is "to be immersed," "losing" one's self into the interaction; it is to work at organising a shared reality; it is to experience "mutual involvement" in a "shared awareness." Such a complex of notions may, Mair proposes, account for how it is that the actual model or state of play delivered by the falling S shapes of melody may have a very oblique relationship to the words actually uttered.

Again, it is the notion of immersion in a shared awareness upon which tones of voice and movement perform direct actions--melody "hard-wired" into the brain, acting directly on awareness, transforming it, that is the key. "Immersion" describes a shared history, participation in a shared awareness, which permits progressive condensation of meanings: less and less comes to mean more and more. Melody, the mechanism, as direct action on awareness, allows for separation of actual, uttered words from the cognitive impact, or effect (meaning).

Mair refers to the research of Basil Bernstein. Bernstein speaks of restricted versus elaborated code encounters. In an elaborated code, meanings must be made explicit; there is narrative. The result is to individuate. In the restricted code encounter, how things are said, rather than what is said, becomes important (Mair 1977:VIII). Meaning become highly condensed. Following on from Bernstein, Mair contends the restricted code encounter--the restricted or condensed interaction--is properly viewed as the vehicle for intimacy (rapport).

Furthermore, it is this type of interaction which is "closer to poetry in the importance of rhythm, timing condensed and implicit meanings" (1977:VII-18)--all the essential elements of Mair's model of communicative interaction.

Mair concludes:

> It is a well-known feature of rapport that what one says and what knows that the other person will know that one meant can be very different, also that a simple sound like a grunt with a fall or rise shape can have meaning within, for example, a family group that would be quote opaque to the outsider (1978:70).

Urion (1978) documents a singular phenomenon in his investigation of control of topic in a bilingual (Cree-English) meeting. He argues that topic is more complex than merely "what is talked about" and proposes a similar mechanism to account for the discrepancy between what is said and what is meant. Topic is not self-definitive. Urion distinguishes between the social context of a statement, and the temporal, sequential context (1978:133). The latter describes "that which is manifested by the statement itself"; the former "that which the statement must necessarily mean" (1978:134). The two are not always equivalent. Where they are not, Urion applies the term "ostensible topic" to the former, and "oblique topic" to the latter. Ostensible topic is "a topic that the lexes describe, which is more or less evident to anyone competent in the language." (1978:134). Oblique topic is "a topic that is tacit, and can be stated in descriptive terms only because of experiential knowledge" (1978:134).

Urion argues that although a physical record of the text (lexes) may adequately describe a "surface sequence" of a topic development, outlining the existence or lack of existence of a surface logic of statement sequence defined by, and explicable with reference to the common-sense meaning of terms, such description is not always adequate to reality. What is really going on may not be what the words tell us. In order to "understand," it is necessary to attribute to interactants a "shared knowledge of an unstated topic sequence, and the logic of its progression as internal to the discourse itself" (1978:137).

Urion maintains that both "ostensible" and "oblique" topics are accounted for by participants in a communicative event. At the same time, he argues, "participants continuously distinguish between the two and reconcile the two in the creation of an event" (1978:196). The oblique topic is in the interactional domain, opposed to a referential (lexical) domain, and each domain is conceived of as a plane, like the face of a Necker cube. Individuals and/or subgroups within a group are thought to "manipulate to direct other participants' attention to one or another of the planes, or to both; to identify a focus that indicates where the import of an utterance-act should be determined." (1978:196). He proposes some specific mechanisms through which definition of topic, as well as regulation of topic change, are achieved. They are (1978:197) "speech (and its counterpart, silence) as an intimate physical event; and meaningful other action, including kinesics." He argues that control of topic may be described in terms of amount talked, code and register shifting, regulation of the sequence of talk, and power--here conceived of as the ability to "perform"; that is, to attract other participants' to attend to one's talk.

Urion's account of the process of communicative interaction approximates Mair's. Both posit a not immediately apprehensible "mechanism" which is thought to regulate the temporal progression of discourse. In Urion's terms, Gunter would seem to have focused analysis exclusively on the surface level of the lexes. Dealing solely with ostensible topic, he is (not surprisingly) able to account for neither the commonly perceived coherence of conversation, nor the often-experiences and intuitively obvious ability of participants to "project" topics at their conclusions in communicative interaction.

4.2.6 Limitations

Although it is an appealing construct, there are some problems with Mair's model of human interaction. Perhaps the most obvious "problem" with Mair's model (at least for those whose interest lies in the study of language) is the exclusive focus on English. If proposed universals of interaction are to be established, Mair's theory must be tested on other languages. There is a need to look at, for example, tonal languages (e.g., Chinese), "fixed stress" languages (e.g., Ukrainian), and "no stress" languages (e.g., French), as well as other languages like English that have variable, semantically related stress.

4.2.7 Rapport

Mair proposes to account for the process of communicative interaction. To date his analysis has tended to focus on interactants in a state he defines as "rapport." Although it is expected that in time the model will be extended, it is of no small import here that Mair has as yet offered no clear definition of the fundamental concept of "rapport." Rapport is, he asserts, "immersion in the here and now"; "a mode of experience into which individual consciousness is lost." It is "something" which interactants can, Mair argues, slip into and out, as it were. He proposes that to be "in rapport" is to move in time physically and cognitively with one's interactive partner(s). Rapport is argued to be an aspect of the broader category of experience which Mair terms "entrancement." Mair connects "rapport" to the notions of synchrony and cognitive rhythm in interaction. He argues that the degree of synchrony in interaction is contingent upon the degree of rapport between interactants. He also argues that, in a state of rapport, the timing of cognitive change point has a systemic relationship to the timing of the physical process of interaction.

But, with all this, what exactly is rapport? It is hard to be sure. Mair seems to be appealing to and relying upon a stereotypical, culture-bound notion. His definition of rapport as "immersion" in the here and now is prescriptive and poetic, but of questionable use in scientific description. The attempt to define rapport through specification of the degree of congruence between the timing of cognitive and physical timing processes is not quantified explicitly and specifically. The vague nature of the notion of "rapport" comes out in Mair's own work: he chose his selected fragments of text for micro analysis from subjective criteria--they were "aesthetically pleasing."

Finally, in his published discussion Mair makes no mention of the "fact" that during interaction interactants may "pretend" rapport, "faking" the outward signs argued to indicate conditions of rapport. He does not deal formally with the problem of pretended rapport or the fact that we can lie.

4.2.8 Projection

The notion of "projection" is equally difficult. Again it seems that Mair asserts his description: "in spontaneous interaction we do in fact project things forward, project topics at their destinations." He suggests a mechanism by which topics are projected at their conclusions--fundamental frequency contour and movement--and proposes that it is their action upon the immediate state of play which in fact projects topics to their outcomes. Mair also claims for interactants an awareness of the projection mechanism such that they can and do predict when in time physical and cognitive models will "come down," but the claim is only tentative until longer speech segments, whole discourses, may by analyzed.

4.2.9 Tropisms

Mair proposes "tropisms" as extremely powerful "final common pathways" in human interaction, much like the "fixed action patterns" which ethologists attribute to birds. There are six tropisms; each is argued to be one "element" of the "force field at now" which is thought to constitute immediate awareness, the shared state of play. The whole notion of "tropisms" is very speculative and highly argumentative. Criticism here, for the present study, might thus be best limited to and circumscribed by Mair's own summary statement: "NB: The identification and definition of tropisms is at present a little haphazard." (1980:24)

4.2.10 Control as Dominance

Mair acknowledges "vulnerability" in the give over of control during interaction. However, although this notion is submitted as part of his model, Mair expands very little on this aspect of the process of interaction. His focus clearly lies on the Machiavellian "melodic competition" and the take-over of interaction: "text control is the outcome." He who controls melody controls text. Mair describes the individual in interaction as part of a supra-individual system in evolution, but then seems to overlook the "fact" of individual awareness in interaction. He leaves himself little or no space to address both (all) individuals involved in any given interaction, rather concentrating on that person who, having accessed control of the melody, apparently controls the text. In Chapple's terms Mair focuses almost exclusively on the "pacemaker" individual in his (Mair's) analysis. This is perhaps the major problem with Mair's model, at least for the present study.

Mair postulates a "social brain," the existence of which has been obscured to us by the fact that the connections between the nodes are not flesh and blood, (i.e., neuronal), but the sensory modalities of the human communication system. The individual is, then, a node in the network of the social brain, organising it and organised by it. Mair's discussion emphasises the latter, almost to the exclusion of the former. He glosses over the fact that for the "social brain" to exist individuals must work to create and maintain it. He does not account in his model for allowance of control in interaction, which I would argue is control. In Mair's model, control is seen as a one-way dominance phenomenon.

The model thus takes on the qualities of an entrainment model. The inadequacy of the entrainment model to the reality of human interaction documented above in the critique of the work of Scollon and Chapple (see Chapter 3) pertains here. Briefly, the entrainment model is a stimulus-response model and is far too simplistic for the "real world." As Chapple points out, there is action without outside stimulus; and action and reaction are not always one-to-one correlates of any given event. In the entrainment model there is no allowance for the individual, for conscious and/or self-generated action or thought; for individual "awareness"; for (existential) choice.

4.3 Mair's Model as a Natural System Model

A complete model of human interaction must account for both individual and system. Neither is complete without the other. (Neither is possible without the other.) We must address the reality of natural systems, taking account of the system as a whole and of the individual "units" which compose the system. In communication systems particularly it is imperative to account for not only the system but for the individuals in and of that system. We must attend to the fact of our ultimate solitude. It is apparently the fate of human beings to live in groups, and die alone. Mair's construct, and seemingly the entirety of research that deals with the process of communicative interaction, lacks an effective means to contextualize itself in such terms, in terms of the natural systems which form part of, create, and in the end are that which they seek to study. What seems to be needed is something that will demand specification of criteria for the choice of domain, unit, and/or level of study. The need is for a methodology that can facilitate definition of "what is to be studied" as well as how "it" is to be studied. Humberto Maturana and Francisco Varela (1975; Varela 1976) propose a theory of observation of natural systems, including human language which may be of use.

Maturana and Varela (1975) argue that the primordial act of indication, the act of separating forms, distinguishing what we are from what we are not, is the basis of every description of the universe. They insist further that "every experimentation and observation implies a theoretical perspective, and no experimentation or observation has significance outside the theoretical framework in which it took place" (1975:12). They deny the "classic" Western paradigm of observation which demands that the observer not enter in the description of that which s/he observes. They argue that in fact the reality of observation and description may be more accurately represented by what Varela calls the "communal" paradigm of description: description shall reveal the properties of the observer; the properties the observer discovers in an observed system depend upon the properties of the observer in relation to the properties of the observed system.

Varela (1976a:64) characterizes description of natural systems from this, the communal perspective:

> . . .for every system there is an environment which can (if we so decide) be looked at as a larger whole where the initial system participates. Since it would be impractical to do this at all times, we

> often chop out our system of interest, and put all the rest in the background as 'environment'.
> . . .To do this on purpose is quite useful; to forget that we did so is quite dangerous.

Varela asserts that in the act of describing natural systems we select from the continuum of the universe, we "chop" out of a "mesh" or "net" an element "A" which we then treat as an initial element at a time zero. Varela emphasises that the act of description requires somebody to do it and sometime in which to do it (1976a:64). The act of chopping, description, specifies a set of relationships among the nodes of the original timeless network, in time. This provides a static representation of the relationships of the system and may be characterized by a "tree."

The tree is a self-recursive closed system, given the interrelatedness of the nodes in a mesh. That is, traversing the levels, the branches, of the tree one ultimately returns to the arbitrarily selected node "A." To get back to the net from the tree we must "eliminate" time. The net is a whole, a set of simultaneous interactions of parts which exhibit stability as a totality. Nets, wholes, exhibit a certain recognizable stability arising from their closure (or organisation). Nets characterize the continuum from which we segment in order to "make sense" of the universe. The parts, the trees, are indications of specific interactions which we can chop out from the whole and consider in detail, as participants in the various processes that constitute the whole. The whole reappears with the perception of the stability of the whole.

The observation of natural systems requires the specification of pairs or dualities, such as "net/tree," in other than hegelian dialectical terms. In classic, hegelian dialectics the notion of dualities describes opposites in conflict. All pairs are of the form "A/not A." They are specifiable only within the same analytic level and are characterised by a basic symmetry. But, Varela argues, there are no such opposites in nature; we need a new dialectic. He proposes cybernetic or post-hegelian dialectics in which "dualities are adequately represented by imbrication of levels"; where one term of the pair "emerges" from the other (1976a:64). The new dialectics is expressed in the form "A," and "processes leading to A," where "A" is a closed system and "processes leading to A" is an open one. In the new dialectics the pairs become complimentary: they mutually specify one another. The notion of duality as opposition and, further, the notion of "duality" itself disappears. The mutual specification of the elements effectively relates them. Varela contends it is possible to observe these dual pairs from a meta-level where they become a "cognitive unity, a second-order whole" (1976a:63). The "logic" behind hegelian pairs is negation. The logic behind Varela's dialectic is self-reference, a notion he argues has been omitted from the theoretical monument of the West, except in very crude forms.

The contextualisation of observation transcends levels and requires the dialectic proposed by Varela with its "trinity" of terms: "A," "processes leading to A," and the relationship between imbricated levels. Varela (1976a:64) explains further:

> Pairs of opposites are, of necessity, in the same level, and stay at the same level for as long as they are taken in opposition and contradiction. Their effective interactions are not (cannot be) specified.
>
> Pairs [in the cybernetic sense] bridge across one level, and this crossing is operational. They mutually specify each other.
>
> It is, of course, the case that when we look to natural systems, nowhere do we find opposition apart from our own projection of values. The pair predator/prey, say does not operate as excluding opposites. Both generate a whole unity, their ecosystemic domain, where there is complementarity, mutual stabilization, and benefits in survival for both., So although we can project values to the opposites predator/prey, the effective unity is a larger one ... ecosystem/species interaction.

Varela describes "reality" as composed of a hierarchy of levels of imbricated systems from less to more complex. But imbrications are observer-imposed specification of analytic levels; their specification must appeal "to a principle which specifies analytic levels with reference to explicit and other-than-arbitrary criterion" (Urion 1978:34). Varela proposes to define a level as "any one step in this ladder" of a hierarchy of systems. He suggests that each level may be characterised by the stability of the system at that point.

The notion of the arbitrary nature of identification of the initial element or node is crucial to the theory of observation of natural systems. It is a problematic notion requiring much elaboration on Varela's part:

> In this primordial act we separate forms which

> appear to us as the world itself. From this starting point, we thus assert the primacy of the role of the observer who draws distinctions wherever he pleases. Thus the distinctions made which engender our world reveal precisely that: the distinctions we make--and these distinctions pertain more to a revelation of where the observer stands than to an intrinsic constitution of the world which appears, by this very mechanism of separation between observer and observed, always elusive. In finding the world as we do, we forget all we did to find it as such, and when we are reminded of it in retracing our steps back to indication, we find little more than a mirror-to-mirror image of ourselves and the world. In contrast to what is commonly assumed, a description, when carefully inspected, reveals the properties of the observer. We, the observers, distinguish ourselves precisely by distinguishing what we apparently are not, the world (1975:22).

4.3.1 Autopoetic and Allopoetic Systems

Maturana and Varela (1975) posit two types of systems in nature: autopoetic systems and allopoetic systems. Autopoetic systems are living systems, such as a human being. An autopoetic system is "an homeostatic (or rather relations-static) system which has its own organization (defining network relations) as the fundamental variable which it maintains constant" (1975:4) The defining characteristic of autopoiesis is thus not components, structure, or function, but organisation. All else, all changes, are subordinated to maintenance of systemic organisation.

Maturana and Varela (1975:6-8) document several "consequences" of autopoetic organisation:

1. Autopoetic systems are autonomous.
2. Autopoetic systems have individuality. That is, by

maintaining their organisation as an invariant through its (the organisation's) continuous production autopoetic systems actively maintain an identity which is independent of their interactions with an observer.

3. Autopoetic systems are unities because, and only because, of their specific organisation; their operations specify their own boundaries in the process of self-production.
4. Autopoetic systems do not have outputs or inputs. They are closed systems which react to perturbations in the medium in which they exist by changing, or not changing, their internal state. This internal "reaction" Maturana and Varela term "compensation." Compensatory changes are subordinated to the maintenance of the organisation of the system at all times. Any relation between these changes and the course of perturbations in the medium pertains to the domain in which the system is observed, not to the organisation of the system.

Allopoetic systems are nonliving systems. Their defining characteristics are roughly opposite to those of autopoiesis. Allopoetic systems are not autonomous. Allopoetic systems have no individuality; they depend on an observer for identity and definition of boundaries. Definition of boundaries is achieved for allopoetic systems through observer specification of input and output surfaces. This specification process defines what pertains to the allopoetic system in its operation.

The human nervous system in this context, i.e. an autopoetic system, is necessarily considered a closed system, changing (or not) its internal state in response to perturbation in the medium in which it exists. Maturana and Varela define the interactions of an autopoetic system with the medium in which it exists as constituting a "structural coupling" with that medium. (The medium, needless to say, may include other autopoetic systems.) The domain in which such interactions occur Maturana and Varela term a "consensual domain."

> The domain of interactions of an autopoetic unity is the domain of all deformations that it may undergo without loss of autopoiesis. Such a domain is determined for each unity by its structure, the particular mode through which its autopoiesis is realized (Maturana and Varela 1975:53).

"Perception" in this framework is equivalent to the organism's, the system's, view of whatever impinges upon it, and this "systemic" view depends upon the "reference" the system is set up to maintain, i.e. its organisation, rather than any externally defined "stimulus" that may be specified by an external "agent." The fundamental difference between the notions "input-output" and "perturbation-compensation" is found in the location of emphasis: the former emphasizes the design of the system; the latter emphasizes the stability of the system (which derives from its organisational closure) (Varela 1976a).

4.3.2 Communicative and Linguistic Domains

Maturana and Varela specify two particular consensual domains of especial import to this discussion: communicative and linguistic domains. The communicative domain is that which allows "behavioral coupling" of autopoetic systems. Communicative interaction is defined as a coupling in which the autopoetic "behavior" of organism A acts as a perturbation for organism B, whose compensatory behavior acts in turn as a perturbation for organism A, whose compensatory behavior acts as a perturbation for organism B, ... and so on until the behavioral coupling is "interrupted" (Maturana and Varela 1975:76). The characteristic of "plasticity" in an autopoetic system allows permanent modifications in the autopoetic system as a result of communicative interaction.

The linguistic domain is defined as "a consensual domain of interactions in which the behaviorally coupled organisms orient each other with modes of behavior whose internal determination has become specified during their coupled ontogenies" (1975:77). That is, the linguistic domain evidences growth, evolution, and expansion as the coupling of the autopoetic systems continues. But, Maturana and Varela insist upon the separation of the linguistic domain and the domain of autopoiesis. They are different, nonintersecting domains even though one generates elements of the other.

4.3.3 The Observer as a Natural System

Maturana and Varela go on to delineate two further "remarkable properties" of autopoetic systems. First they make a definitional statement: an autopoetic system with a nervous system is capable of interaction with its own states, capable of developing with others a linguistic consensual domain, and can treat its own linguistic states as a source of deformations/perturbations (1975:77). An autopoetic system can therefore interact linguistically in a closed linguistic domain. "Through recursive interactions with its own linguistic states it [the autopoetic system] may always remain in a position to interact with the representations ... of its interactions" (1975:78). The "remarkable property" describes a system commonly known as an observer. Maturana and Varela argue that this domain of recursive interactions, the observer domain, is, in principle, infinite: "there is no moment when the system will not be in a position of recursively interacting its own states, unless autopoiesis is lost" (1975:78).

The second "remarkable property" refers to the notion of self-observation. A living system which can be an observer can also generate a domain within which it is an observer of itself as an observer. This process occurs through interaction with those of its own descriptive states which are linguistic descriptions of itself (1975:79). Maturana and Varela define this state as property of the domain of selfobservation, and further define self-conscious behavior as self-observing behavior (i.e. behavior in the domain of self-observation). They emphasize that the observer as observer remains forever in a descriptive domain, forever locked into a relative cognitive domain. From this domain no description of an absolute reality is possible. Any representation of even a presumed interaction with the "absolute" would be tempered, irrevocable altered, and determined by the autopoetic organisation of the observer. The cognitive reality, the representation "abstracted" from experience must, by definition, be relative to the knower.

Varela (1976a:65) ascribes three main characteristics to an observer:

1. capacity for indication: to decide boundaries, to come up with nodes, systems to have criteria for stability
2. capacity for time: to chop a net and sequence, to compute through a process, to approximate the stability of a whole
3. capacity for agreement: to externalize, to synchronize with other observers, to re-produce other's distinctions and follow corresponding time-patterns.

He argues that "what happens" with, between, an observer and a system is appropriately a property of their structural coupling. "What happens" is system dependent in that the organisation of the system will determine how it (the system) "behaves" during coupling. "What happens" is observer dependent in that the organisation of the observer must determine his behavior during coupling; that is, the observer's autopoetic organisation will determine what perturbations may be introduced at the system under observation. An observer interacts with an (other) autopoetic system in a consensual domain which the two have mutually specified. This means the observer must in his observation specify a domain of the internal state of the observed organism nonintersecting with a domain of interaction, and nonintersection with the domain from which the observation is made. The domain of the observer is always a descriptive domain. Notions such as ontogeny, development, function, and purpose, form part of the observer's domain alone.

> [They] have no explanatory value in the phenomenological domain which they pretend to illuminate, because they do not participate as causal elements in the reformulation of any of its phenomenon. This does not preclude their being adequate for the orientation of the listener towards a given domain (Maturana and Varela 1975:15).

Observer and observed are mutually revealing. Whatever is seen, heard, understood, or described belongs to the consensual, descriptive domain which is itself a reflection of the properties, the organisation, of the observer and the observed.

4.3.4 Implication of the Theory of Observation of Natural Systems

The theory of observation of natural systems proposed by Maturana and Varela has major implications for the study of the process of human communicative interaction, and the study of the temporal structure of speech. First, participants to a communicative interaction receive new definition: participants are necessarily autopoetic systems, describable primarily in terms of organisational closure. The are individuals engaged in a process of structural coupling with other individuals. The result of the coupling may produce an autopoetic system of a higher order, "society."[18] If produced, the "higher order autopoetic system" must have definite organisational characteristics and properties of its own, but the individuals involved in the creation of the system never surrender their autonomy, their autopoiesis. By definition they never can, or they and the system will cease to exist. In addition, each individual interacting in the system is necessarily a self-observer.

Second, the role of participant-observer is given new definition. The observer is brought to the foreground. The observer-researcher is required to admit and specifically, explicitly acknowledge participation in the observed system, even while posing as a "super-observer" of the entire system. At the same time, observations are contextualized, placed squarely in the domain of the observer. Observations, as properties of a descriptive domain, pertain only to that domain and manipulations of properties within that domain. They have no explanatory value outside it: descriptions are not causally related to the actions or inactions of autopoetic systems involved in the consensual domain. The theory of observation of natural systems demands that context be recognized and made explicit. In this manner Maturana and Varela may be seen to demonstrate that studies which artificially deny context, whether it is social or biological, must explain phenomena at a "trivial" level (Urion 1980:513).

---

[18] On this matter Maturana and Varela make little comment other than to indicate disagreement among the authors.

Third, Urion (1980) notes that the notion of recursion--observing that one observes--is central to the concept of self-conscious operation in linguistic and communicative domains. They are both recursive consensual domains. The first consensual domain arises in the initial act of indication, the arbitrary separation and segmentation of continuous phenomena. The initial act of indication establishes the "paradigm" within which the autopoetic system is said to operate. That is, the initial act of indication establishes paradigmatic relations. Urion proposes that description of syntagmatic relations must imply "multiple-order recursions of observation" (1980:513); and further, that the differential levels of recursion required for paradigmatic and syntagmatic relations specify "a context for discussions of both uniformity and variability in language behavior" (1980:514). What is more, as Urion (1980) argues, the notion of self-observation in a consensual domain is powerful in its prediction of the aspects of "novelty and creativity" in communicative interaction.

4.3.5 Descriptive Adequacy of the Theory

The descriptive adequacy of the model proposed by Maturana and Varela is evident in its applicability. The work of Scollon, Chapple, and Mair, among others, is readily translatable into the terms of the theory of observation of natural systems.

Scollon's work fits nicely into the theory. His discussion of tempo and the "temporal bond of ensemble," as well as his notions of "entrainment" and his insistence on (and search for) the biological foundations of the temporal structure in language would seem to be usefully contextualized in terms of autopoiesis and consensual domains in which interactants interact.

Chapple's work is equally translatable. The notion of perturbation in a closed nervous system reformulates Chapple's notion of the "pacemaker individual" in much the same manner as Mair's notion of control: there is manipulation in a consensual domain that allows for individual existence and predicts the ascendancy of allowance of control. Chapple's discussion of entrainment and crowd behavior are likewise susceptible to the constraints of autopoiesis. Varela's notion of "imbrication of levels" contextualizes Chapple's discussion of "polarization" and "polarized arrays." The notion of structural coupling in conjunction with the concept of mutual specification of "dualities" or "pairs" across levels circumscribes Chapple's notion of group or set events and the hierarchical systems in which they are found.

The theory proposed by Maturana and Varela is particularly translatable into the framework put forward by Mair. It is also a particularly useful translation, serving to contextualize Mair's model in terms of the natural systems of the human organism--of these, communicative interaction is but one. Mair's notion of state of play in immediate awareness may translate into Maturana and Varela's "consensual domain." Thus, as Urion (1980:514) notes, "the notion of perturbation in the closed nervous system provides for a process whereby individuals effect topic change by manipulation in the consensual domain." Urion observes further that Mair's definition of rapport fits with Maturana's and Varela's definition of recursion in the observational domain. That is, "[t]he moment of 'rapport' implies a conflation of recursive orders" (1980:514). Mair's notions of "shared awareness" and "shared plan" are also effectively reformulated in terms of the manipulable properties of the consensual domain.

The theory of observation of natural systems provides for a new address to the concept of control and allowance of control in interaction. In a consensual domain, "control" is a manipulable variable, attributable to a text organiser only so long as there is no interference in the autopoetic organisation of the interactants. Mair's assertion--"he who controls melody controls text"-- is necessarily and usefully modified with the notion of closed autopoetic systems "interacting" in a consensual domain.

Finally, the aspect of history and growth, or expansion, which Maturana and Varela attach to their communicative and linguistic domains seems to address the same kind of phenomena of which Mair speaks in his discussion of the progressive condensation of meaning that occurs during the history of an interaction.

4.3.6 Summary

Maturana and Varela have initiated a powerful descriptive device. They provide a meta-language for the description of natural systems. Their theory demands that researchers contextualize investigation in terms of the autopoetic systems which produce the phenomena studied. The theory of observation of natural systems proposed may, in that it allows specification of criteria for the definition of units, enable researchers to "get at" the biological basis, (or the natural science variables, whatever one's choice of term) of their study--in this case the process of communicative interaction, the temporal structure of speech.

Chapter 5

Methodology

5.1 Initial Trials

The Group Processes Laboratory (GPL), made available to me through the cooperation of the Department of Educational Psychology, under whose aegis the room is operated, was used for recording throughout the study. The laboratory proper is a rectangular room. Three sides of the room are in fact one-way mirrors. Behind two of the three walls are observation rooms; behind the third a large control room. There are curtains along all three mirrored walls. The curtains may be drawn to prevent unintended or unwanted observation. The room is carpeted. Furniture varies, but usually includes three or four tables and several wooden chairs that are scattered throughout the room. During taping the two observation rooms were completely curtained off. Only the control room wall was left unobstructed. Tables and chairs were removed to the sides of the laboratory.

There are four black and white video cameras, model GSC CTC 5000, in the laboratory proper, one in each corner of the room. All cameras are suspended from the ceiling. Two cameras are fixed and immobile; two are adjustable. The fixed cameras come with a Canon 12.5 mm CF-12-5-A wide angle lens. Adjustable cameras are equipped with Canon V10X15R F2.8 zoom lens and Cunningham LTS-122 and Pelco CM-101-C pan-tilt mechanisms. Fixed and adjustable cameras hang in opposite corners of the laboratory: fixed opposite fixed; adjustable opposite adjustable. Adjustable cameras may be manipulated--moved vertically and horizontally, and focused--from within the control room. The signal from each camera feeds into a panel in the control room where it is displayed on 9-inch Electrohome monitors. The final signal to be recorded is displayed on 14-inch Electrohome monitors, model V06-219. Control panel capabilities limit simultaneous onscreen display to two video signals.

The two adjustable cameras were used exclusively in this study. The flexibility they offer allowed finer image resolution and appropriate framing of the subjects on the video screen during interaction. A single table and two chairs were placed in the centre of the room, the table angled. Through trial and error, table and chairs and cameras were manipulated so that images of comparable size and quality for each interactant came through on the video monitors. A "split-screen" effect, where both subjects appear facing the camera on screen, was generated with the switcher, Sony model SEG-2.

The control panel also contains an Altec amplifier, audio mixer, and speakers. These allow monitoring of signals from the four microphones, Shure Model 571, suspended approximately 15 cm from the ceiling in each quadrant of the laboratory. The control panel is capable of mixing all and/or any combination down to a single audio channel on a video cassette recorder. A fifth microphone, installed expressly for this study, may be mixed with the other four, or recorded on its own.

Interactions were recorded on 3/4-inch 60-minute Sony video cassettes, KC-60, using a Sony VO 1800 video cassette recorder. For all the sessions but the first, an additional simultaneous audio recording was made. Two high quality Sony 160 microphones were positioned as near as possible to interactants. A tray containing refreshments (coffee, tea, etc.) occupied the centre of the table, between the subjects. Audio signals were recorded with a Sony TC 630 reel-to-reel stereo tape recorder on Scotch 176 Low Noise tape at 19 cm/second. The audio and video records are supplemented by a notations of what I thought of as significant action. Audio and video monitors in the control room and the one-way mirror systems were used to identify and clarify off-camera and/or blurred, ambiguous sound and movement.

The first session recorded for systematic analysis had as subjects two males from New Zealand, M and J. The session was an experiment designed to evaluate the suitability of the GPL and determine the ability the technology available there to produce the kind of data necessary for this sort of micro analysis of human interaction. There was no attempt to alter the laboratory. A table and two chairs were positioned in the centre of the room as noted above. Coffee was provided.

M and J arrived at the university in the evening, after work. They had never been on campus before and were curious. I led a brief tour as we chatted. The tour finished with a quick turn through the department and an "orientation seminar" explaining what exactly was expected of them. They were reminded that they were simply expected to "chat". Immediately prior to taping they were shown through the GPL itself. M and J are long-time friends. As a former journalist, M was particularly interested in the place and role of words, both spoken and written, in the process of human communicative interaction.

The session lasted well over an hour, continuing long after the tape had run out. I remained in the laboratory with M and J for a major portion of the taping session, absenting myself only twice. Both absences occurred towards the middle and end of the interaction and were for substantial periods of time. One absence was required to check on the recording equipment and to ascertain that it was functioning properly; the other was a break that I prolonged in an effort to remove myself from the laboratory for an extended period and allow the subjects to interact on their own, out from under the eyes of a participant-observer.

After the taping M and J asked if they might be allowed to see part of the video tape. We adjourned to the control room where we viewed randomly selected portions of the recorded interaction. A spontaneous discussion of the interaction as data source ensued. M and J's comments ranged from apologies for fatigue to queries about the suitability of the data for the present study, and concomitant questions on the purpose and place of audio and visual data--together and separately--in the study. Both M and J remarked strongly upon the "effect" of the room. They observed that they found the GPL to be an extremely sterile environment and argued that that sterility contributed negatively to the creation of a "good conversation". They felt that in more amenable surroundings, under better circumstances, they were capable of "better" interaction; i.e. one more representative of the actual day-to-day interactions which formed and cemented their friendship, and thus, it might be contended, more suitable for analysis, if a goal of the study was investigation of spontaneous, natural conversation. In the end, though, the overall consensus was that it had been "fun" and that we should do it again soon, with or without the recording equipment.

Several changes were made as a direct result of this session. Analysis of the recorded data revealed several flaws and shortcomings. I was not satisfied with the quality of either the audio or video recordings. Cameras were re-positioned and re-focused, and the location of the subjects in the laboratory was shifted to account for the variable quality of available cameras and to allow for better image resolution on the screen. The position of ceiling microphones was altered in an effort to improve sound quality. All microphones were lowered to a height 50 cm below the ceiling, and an additional microphone was secured directly above interactants. In addition, the extremely poor quality of the audio track on the video cassette demonstrated the need for a separate audio recording device. The audio recording capabilities of the GPL were accordingly supplemented with a reel to reel stereo tape recorder and two high quality microphones. A separate recording track was established for each interactant.

Following the suggestions put forward by M and J, The GPL was redecorated for subsequent series of tapings. Seminar room style furniture was removed. Stuffed and upholstered chairs were brought in. Several large plants were placed strategically throughout the room. Coffee tables were substituted for the higher office-type ones. Audio microphones in table stands were positioned on coffee tables beside each interactant's chair, as near as possible to the interactants. Incandescent table lamps were set on tables. (They were originally intended to replace the harsher overhead fluorescent lights which M and J had found bothersome, but they were quickly reduced to the status of decoration when it was found they could not provide sufficient light for video recording with the equipment available in the GPL.) It was as pleasant a room as I could make it. Finally borrowing from Mair, it was decided that coffee and coffee liqueur were to be provided for subjects during recording sessions.

This was the arrangement used for a second taping with M and J, as well as a session involving another pair, 'C' and 'D', two subjects of the present study. Taping procedure followed established protocols. The actual recording was preceded by an orientation seminar in the GPL, and was followed by a viewing of randomly selected portions of the video-taped interaction, during which there was a general discussion of the interaction as data. Again, I remained in the laboratory with subjects for a major portion of the recording sessions, leaving only to check on equipment or to purposefully absent myself from the interaction arena for a while.

Observation and analysis of data from these sessions led to further changes.

First, the claims made by M and J about the effect of room decor proved, in this case, to be unfounded. Interaction was not noticeably different or better in the redecorated GPL. There was no appreciable difference in the interactions, qualitatively or quantitatively, along any of the descriptive, observational, or analytic axes used in this study. The "charm" of the subjects and a pointed lack of hidden camera techniques seemed to far outweigh and over-ride the effect of the presence or absence of some plants, coffee tables, lamps, and chairs. As it turned out the redecoration was a needless expenditure of time and energy and was discontinued in further tapings. Coincidentally, while examining the recorded data from these sessions I had decided to limit description and analysis in the visual mode to movements of the head and hands. The original arrangement of table and chairs was particularly suited to this. The table around which the subjects sat while conversing was a convenient and absolute cut-off point; it furnished a base-line for observation of movement and, in this way, helped frame the study.

Second, the provision of alcohol was problematic. Alcohol was an unknown, potentially confounding variable. It did not seem especially to enhance the quality of the interaction; nor did it seem to especially harm it, but it remained an unaccounted-for variable. As a result, a decision was made to eliminate alcohol from the refreshment list for all future recording sessions in the study.

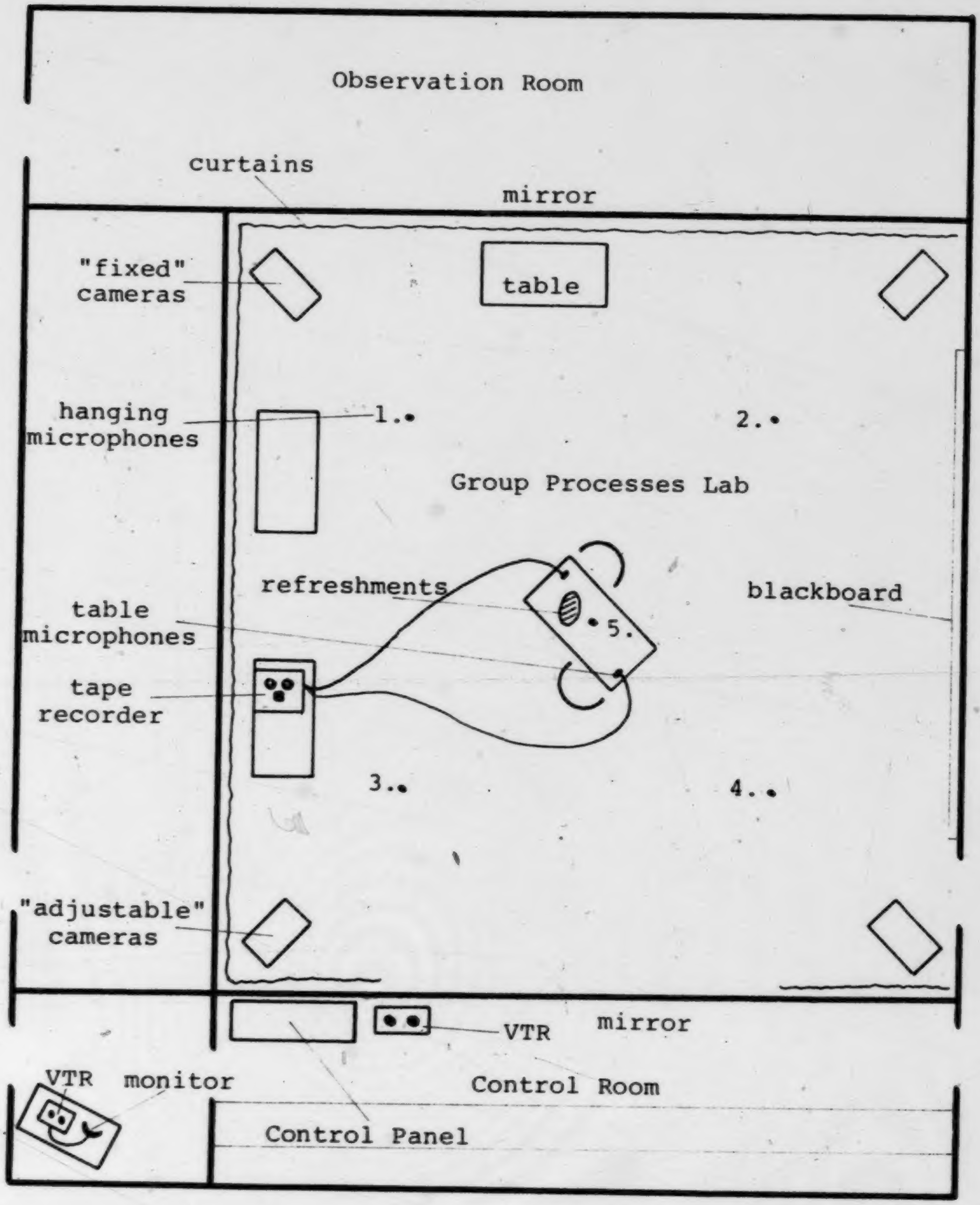
Third, the continued presence of the experimenter in the laboratory with the interactants was proving difficult. On a very basic level, it was impossible to remain in the laboratory and to manipulate and to check on recording equipment properly during the taping. The result was often a recording of inferior quality and of no value to the study. What is more, it was found that the researcher's presence in the laboratory during recording was apparently not a necessary facilitator for the creation of rapport and "satisfactory" interaction. Interaction carried on with or without me. And, on those occasions where I was brought into the conversation by one or another of the subjects, the text became a "trialogue," with one participant perpetually off-camera. During these episodes there is continued evidence of what is here interpreted as synchrony, but the data are unavoidably and irreversibly skewed, carrying the project necessarily into investigation of triadic interaction. However interesting, such research is beyond the scope of the present study. Consequently a decision was taken to remove the researcher completely from the laboratory during recording. I would remain in the control room for the duration; from there I could properly monitor the recording process and, when desired or necessary, observe the interaction first hand through the one-way glass.

The layout of the laboratory for all subsequent recording sessions is represented diagrammatically in Figure 5-1.

5.2 Selection of Subjects

Mair had selected a single pair, a male and a female from the same freshman class, as subjects for his primary data base. I was interested in providing a broader perspective from which to evaluate the theory. First inclinations were to record pairs of speakers in several languages, but financial and temporal constraints proved prohibitive. Translation alone was an impossible imperative. Second inclinations required pairs of subjects specified by gender. The purpose of the study was not to investigate "kinds" of synchrony, attaching "emotion" words to instances

Figure 5.1
The Group Processes Laboratory


Observation Room
curtains
mirror
"fixed"
cameras
table
hanging
microphones
1.
2.
Group Processes Lab
refreshments
blackboard
table
microphones
5.
tape
recorder
3.
4.
"adjustable"
cameras
VTR
mirror
VTR
monitor
Control Room
Control Panel

of what were taken to be synchrony, but rather, to establish the existence of synchronous behaviours in as large a context as possible. It was thus decided to record in English three separate interactions: male-male; male-female; female-female. There were to be four subjects: two males and two females. A central male-female pair was to be recorded first. They were to then each choose a second partner of the same (their own) gender for a second taping session.

Subjects were evaluated for selection along several axes. To begin with the obvious, subjects had to be willing to talk "on tape." In addition, of those so willing, only those with a "minimal" knowledge of the study were considered as possible candidates for the study. That is, no invitation was extended to those involved in the development and progression of the work. Subjects were simply to "chat," aware of the general purpose of their interaction, but innocent of the details of the specific purpose of the study. Subjects were chosen from among graduate students in the Department of Educational Foundations. The decision to confine selection to those graduate students in fact constitutes a kind of control for at least four variables of import to research studies involving human subjects: age, socio-economic status, level of education, and background experience. Subsequent choice of subjects was in some ways a random selection of graduate students in the department.

In the end the subjects selected were chosen because:

1. I knew all the subjects personally, and they all knew

one another and were on friendly and congenial terms;

2. they had expressed an interest in my work, and had readily agreed to be recorded on audio and video tape;
3. they were familiar with the technology of audio visual media and were not likely to be intimidated by the presence of cameras, microphones, and tape recorders;
4. they all had experience as teachers both in the school system and at the university, were comfortable in front of groups, and were accustomed to being watched closely while speaking.

5.3 Subjects

D was the first subject chosen. He was a frequent coffee companion. Over time he had expressed an interest in the study and offered himself as subject. I asked him to suggest a female and a male in the department with whom he would be willing to talk for at least half an hour, "on tape." D mentioned three or four suitable candidates, C and E among them. C was approached first. C and I had taken a course together the previous semester and, in discussions of our respective thesis topics, she too had volunteered herself as subject. C readily agreed to participate in the study and suggested L, her office partner, as the female with whom she would prefer to converse for the recording session. When approached, both E and L freely consented to take part in the study. They knew very little about my work, but indicated that what they knew they found intriguing.

They were especially concerned with ramifications of the findings of the study vis-a-vis communicative interaction in the university and public school classroom.

D and E were part of the sociology group in the Department of Educational Foundations. They had taken courses together and were often coffee mates. At the time of the taping both D and E were well into their respective Ph.D. programs, but E had taken the year off to take part in the Intern Program offered by the Department of Education, Government of Alberta. They had not seen each other for at least a couple of weeks.

C and L shared an office for one school year or more. They had taken courses together. C and L were part of the History Group in the Department. At that time they were both Master's students, more than half way through their respective programs. They too, had not seen each other for some time; conflicting schedules had minimized contact. Both indicated to me that they were looking forward to the taping session as an opportunity to "catch up" on one another.''

C and D had adjacent offices. They were occasional coffee mates and seemed to enjoy entering into long, complicated, and very often rather boisterous discussion-arguments.

''In fact, from the moment they saw each other until well after the tapes had run out, C and L dove in and refused to surface until they had said what they needed to say. On our way to the GPL, for example, they ignored my attempts to put them at ease. Once there they paid me no heed, remaining oblivious as I puttered about the laboratory performing the customary last minute pre-recording checks and adjustments.

D is a large man, well over six feet in height. He was an all-star tackle on a Canadian college champion team. He is from the east; he grew up in Montreal, and took his initial degree from McGill University. Before coming to Edmonton, he completed his Master's degree at the Ontario Institute for Studies in Education, and taught for two years at an inner city technical high school in Toronto.

E came to Edmonton from Winnipeg, Manitoba, where he spent much of his youth. In the year prior to entering the Ph.D. program at the University of Alberta, E had completed his Master's degree at the University of Manitoba, had worked in the Winnipeg Public School system as a full-time kindergarten teacher, and had lectured part-time at the University of Manitoba.

C is, by her own description, short. She has long dark hair. She is an Albertan, born and raised on a farm in southern Alberta. She has extensive experience as a public school teacher, both in the city and in small town Alberta.

L is the youngest subject in the study by about four years. She is of average height, has blond hair and blue eyes, and wears glasses. L came to the Master's program directly from her undergraduate studies in the Faculty of Education. Her teaching experience has been mainly in university where she has taught a first year course in the history of education in Canada for about two years.

5.4 Protocols for Analysis

The object of analysis of the recorded data was to provide as full an accounting of the factors which describe the process of human interaction, to attempt description of interaction in terms of the "artificially extracted parameters" of Mair's model, and to isolate and describe instances, moments and passages, where smooth, finely detailed sequencing and/or simultaneous manifestation of sound and movement might be said to constitute moments, instances, passages of what is here termed "synchrony". To that end the data were examined and evaluated in the following manner.

The entire corpus was examined aurally and a transcription in standard notation was made of the entirety of each half hour conversation from the audio tapes. A time code, accurate to one thirtieth of a second and indicating frame-by-frame advance was "burned onto" the video tape recording of each interaction. It occupies a good portion of the lower quarter of the video screen. A grid of lines, interspersed with dots--line scale = 2 cm; dot scale = 1 cm (on a 17 inch video screen)--was also incorporated directly into the video signal. The parallax problems which plagued Mair's study were thereby eliminated. Finally, time code, grid, and interaction record were dubbed from three quarter inch video cassette to half inch open reel. This latter process was necessitated by the accessibility of available technology. The study requires time-consuming, slow-motion, frame-by-frame analysis of selected fragments of text. The only available playback machine with slow-motion, frame-byframe advance capabilities was half inch open reel Panasonic video recorder and playback unit, model number NV 3130. Through the kind cooperation of the Instructional Technology Centre, Faculty of Education, I was able to obtain free and continued access to it as well as access to a seventeen inch Sony Trinitron television monitor. For the purposes of this study, the playback mode of the Panasonic was altered slightly to continue to provide sound playback in the slow-motion mode. This allows for more precise location of vocalisation (voice onset and offset) on the tape.

It was with this array of data that the study began in earnest. I watched and listened to the audio and video records again, and again, and again in the hope that some way to enter description might, with sufficient exposure, become obvious to me; that some aspect of Mair's model would, in time, manifest itself as "primary and obvious," the place from which to begin. The evaluative metric of each of the components of his model was the power of the construct to describe the complex of recorded sound and movement which constitutes the progression of observable events in interaction. I examined the audio track with the video, the video track without the audio; I ran the video tape at full speed, in slow-motion; I made a series of notes from sessions with the recorded data, using the descriptors provided in Mair's model, and compared them for adequacy and accuracy. It was in such a way that Mair's final "artificially extracted parameter," the notion of synchrony, emerged as ascendant in this context. Here was my "primary and obvious" aspect of the model.

In his data, in his video tapes and laryngograph traces, Mair observed "something," a phenomenon he called synchrony. The result is a set of "mutually dependent notions" which makes unique, perhaps extravagant, claims about the process of human interaction and the nature of the reality in which it occurs and which it creates. The organisation of the data of the present study, though similar to Mair's, are not identical. The differences in technologies available by themselves prohibit identity. Nevertheless, there is something in the data similar to that recurring phenomenon described by Mair as synchrony. This study, then, endeavours to substantiate the claim that synchrony does in fact exist and is identifiable and observable in the tapes. There is no attempt to validate the entire model; that is not the present concern. The aim is to explore Mair's observations about synchrony, not the claims he makes to explain those observations.

Several axes were explored for the identification, definition, and documentation of synchrony, as defined by Mair, in the data. Sound and movement, manifestations in the aural and visual modes, emerged as the most salient descriptors. Within an individual, simultaneous occurrence in the two modes, and/or production of speech and movement in temporal sequence too close to allow for the intercession of a stimulus-response "feedback" loop is described as evidence of intra personal synchrony. Between individuals, such simultaneous or closely sequenced temporal eventuation constitutes inter personal synchrony.

The entire corpus was examined several times, over several hundred hours, to find selected bits of interaction for detailed transcription that was to document the existence of synchrony in its several and various forms. The magnitude of the data sample posed an immediate problem. There were "synchronies of all sorts" in evidence throughout the corpus. How was the corpus of data for analysis to be limited to a manageable size? A decision was taken to search for a single, "obvious" example from each recorded interaction. Each example chosen would be of extended duration, i.e., longer than three seconds, and would evidence throughout its duration simultaneous manifestation of sound and movement with and between interactants including "coming in" together after silence, and, the close temporal involvement required for the smooth sequencing of repartee. Fragments were selected for micro analysis on the basis of integrity. That is, selected fragments evidenced a continuum of instances of synchrony that was conspicuous in its inclusion of all aspects of the concept. The multitude of fragments excluded from the study, many of which are spectacular, were found to be lacking along any one or more of these axes of distinction.

Initial selection was admittedly subjective--bits were chosen because they were aesthetically pleasing; they looked and sounded good. My initial selection of fragments, five from each tape, was checked and validated by two researchers in interaction analysis, both individually and conjointly. Subsequent verification and evaluation was quantitative, in terms of the parameters of Mair's model, as well as qualitative. In this manner the single, "obvious" example from each recorded interaction was chosen for micro analysis.

For each fragment there are five measurements: fundamental frequency/basal intonation contour, voice onset and offset, and manual frame-by-frame plotting of face (head) and hand movements. The latter three are the visual variables. The visual record of head and hand movement is obtained by monitoring a single spot on the face and a single spot on each hand as they change position between successive single frames on the video tape. The slow-motion mode of the Panasonic video recorder delivered stable single frames for plotting. Individual frames were identified, defined, and sequentially numbered in time by the time code. Following Mair, the interval between frames was chosen as one tenth of a second, because, as he argues, inertia of the head makes notable changes of direction "impossible" within a shorter time period, and, furthermore, perceptual critical fusion frequency is around this speed and makes perception of faster changes unlikely. For a single frame selected face- and hand-points were established by noting vertical and horizontal coordinates (x,y) in a two-dimensional space (the screen, and then the graph paper). Time is included by following these points across successive frames on the video tape.

Fundamental frequency contour and voice onset and offset, the acoustic variables, were obtained by the use of a mingograph, trans pitchmeter, pulse generator, oscilloscope, tape recorder, graphic equalizer, and cassette recorder. The mingograph used in this study, the Mingograf 34, is a four-channel ink-jet "strip" recorder. It delivers a continuous "hard copy" of an input signal. The mingograph is generally regarded as more sophisticated than most mechanical recorders, and has a higher frequency response (up to approximately 650 Hz). The ink-jet pens, in fine line of ink onto the paper, deflect back and forth as the input signal frequency rises and falls, converting the electronic signal to a graphic plot. Paper speed may be varied from 2.5 to 100 mm/second. Paper speed for the recordings of interest was 100 mm/second.

The Trans Pitchmeter takes as its input a recorded acoustic signal and delivers two outputs: a signal proportional to the instantaneous fundamental frequency of the voiced segments of the acoustic signal, and a special signal called a "duplex oscillogram" which may be used to aid segmentation of an acoustic signal into "phonemes." This latter capability was not required for this study. The pitchmeter removes the higher harmonic overtones of the input signal with a low-pass filter so that the output signal is just (all and only) the fundamental frequency representation. The low-pass filter is adjustable, with successive rising filter frequencies of 90, 120, 150, 200, and 300 Hz, to accommodate the variability of the fundamental in different human speakers. The low-pass filter must be set with each data fragment to pass only the fundamental of the waveform being analyzed. The pitchmeter instruction manual suggests some "normal" settings for "types" of data--e.g. 120 Hz and 150 Hz for male voices; 150 Hz and 200 Hz for female voices-but cautions that the "correct" setting for any particular data sample, indicated by a smooth fundamental frequency contour with no sudden jumps or discontinuities, can be determined only by trial and error. The pitchmeter also has a high-pass filter, usually set to 70 Hz, which is used to suppress background hum of 60 Hz frequency. Output of the Trans Pitchmeter is, then, a voltage which is (according to the manual) proportional to the instantaneous fundamental frequency of the acoustic signal. It is this signal which for the present study was fed into the mingograph to be recorded graphically.

Frequency corresponding to pitchmeter output recorded in the mingograph trace was determined through a calibration procedure. The trans pitchmeter is capable of producing 19 fixed tones, the frequencies of which are 60, 70, 80, 90, 100, 120, 140, 160, 180, 200, 225, 250, 275, 300, 325, 350, 375, 400, and 450 Hz, that when recorded on the mingograph serve to form the basis from which a "frequency ruler" is constructed. From this it is possible to measure with reasonable accuracy the frequency of the fundamental at any given point in time.

The Trans Pitchmeter has one further capability of import here: it is designed to "spread out" the part of the frequency scale used for any speaker to "get at the best exactness of the measurements" within the range of amplitudes expressible on the mingograph. (The maximum amplitude of the mingograph trace is approximately 40 mm.) Thus the "zero line" knob: when this knob is pushed in, the pitchmeter expands the scale of the frequencies from 0 to 200 Hz; when it is pulled out, the scale from 150 to 350 Hz is similarly expanded. The former setting usually allows for more exact measurement of the fundamental frequency of male voices, the latter for female voices.

The data of interest involved vocal production of both males and females, sequentially and in concert. Through trial and error it was found that each combination of Trans Pitchmeter settings produced slight variations in the fundamental frequency trace delivered by the mingograph. No one trace proved to be more "correct" than any other. Each trace provided different information, a varying accentuation of aspect, which was invaluable in the complex process of disambiguating confused sections of the mingograph record. For each fragment of text, then, there are six separate mingograph traces, one from each constellation settings: (1) zero line = male; low-pass filter = 120 Hz; (2) zero line = male; low-pass filter = 150 Hz; (3) zero line = male; low-pass filter = 200 Hz; (4) zero line = female; low-pass filter = 120 Hz; (5) zero line = female; low-pass filter = 150 Hz; (6) zero line = female; low-pass filter = 200 Hz; Those traces found to be of particular use in reading the fundamental frequency of each selected fragment are included in Chapter V with the transcription of head and hand movement of each fragment.

The pulse generator, HP Model 8003A, delivers a precisely controlled signal of variable shape and time form. For this study, the pulse generator was set to produce a continuous 10 Hz square signal. This signal was fed into a second channel of the mingograph and recorded in time with the fundamental frequency. It is an accurate measure of time and provides for division of the mingograph representation of fundamental frequency into precise tenth of a second intervals. Elements of both aural and visual modes can thus be examined every tenth of a second.

The oscilloscope, HP Model 1201A, served to check the temporal accuracy of the square wave signal generated by the pulse generator before and after recording by the mingograph.

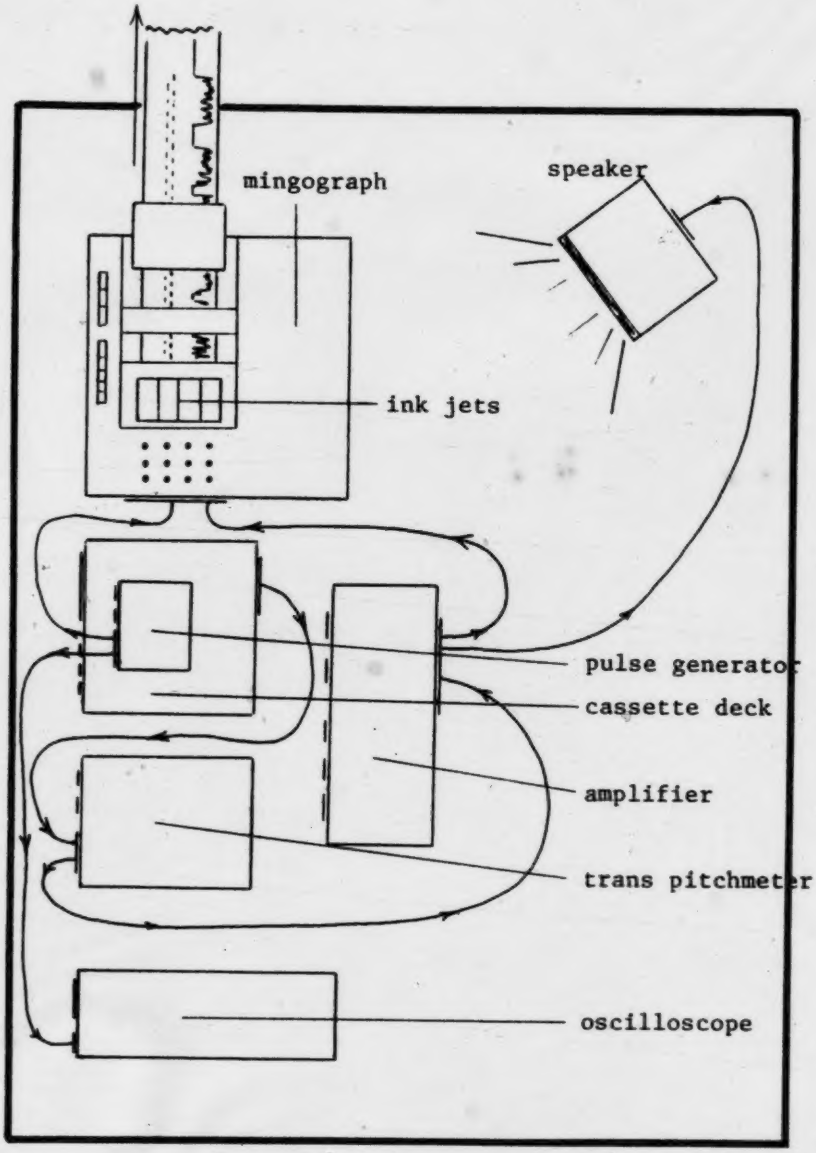
The set-up of equipment that in the end delivered the mingograph traces of fundamental frequency contour may be represented diagrammatically in Figure 5.2.

First trials with the mingograph revealed a surfeit of extraneous noise on the audio tapes which interferred with the fundamental frequency to such an extent that the mingograph trace was unreadable. Even "silence" clear to the ear, for example, was not discernable on the mingograph trace. A graphic equalizer was brought in to clean up the audio recordings. The tapes were dubbed from the open reel, through the graphic equalizer, the ADC Sound Shaper 110, to a cassette. Frequencies around the fundamental--those from 125 to 250 Hz-- were emphasized. All others both above and below were suppressed. The clean tape provided the signal from which the mingograph extracted the hard copy of fundamental frequency used in the study.

Selected fragments were transcribed in standard notation, and the transcription placed in time under the mingograph trace. Placement in time is accurate to within 0.1 second. Placement was achieved first by ear from the video tape and substantiated by fundamental frequency traces.

Selected fragments were also transcribed rhythmically using standard Western musical notation. Syllables heard as "stressed" on the audio recording were marked on a written transcription. The audio tape was then slowed to half speed and syllables heard as "stressed" at 9.5 cm./second were

Figure 5.2
Equipment Set-Up For Mingograph Record


mingograph
speaker
ink jets
pulse generator
cassette deck
amplifier
trans pitchmeter
oscilloscope

marked on a second transcription. These two transcriptions were compared and superimposed. They agreed almost invariably. Next I listened for some sort of regular metric pattern of stressed syllables. For the most part this process was accomplished with the tape running at half speed. I found it easier to determine an underlying metre with the tape moving at half speed. At half speed everything happens so much more slowly, and, as Scollon (1981a:10) argues, this process "distorts the voices enough that one doesn't hear them as the original participants but at the same time does not lose the phonetic detail necessary to hear them as ordinary speakers."

There was no attempt made to impose a "standard measure" on the data. What repeatedly emerged however was a measure in duple metre containing a strong (down) beat, and a weak (up) beat. The selected fragments were accordingly divided into measures of $\frac{2}{4}$ time, where a quarter note represents on pulse or beat, and a detailed rhythmic transcription was completed.

This transcription of notes is admittedly subjective and open to attack as impressionistic. Scollon (1981a:11) summarizes the central problem: "The central problem is to decide whether or not one's hearing of the meter is some form of superimposing one's expected metric pulses on whatever syllable they happened to fall." That is, is the metric pulse heard by observers really there, or, are we (they) putting it there? It is a serious question, a difficult problem, one for which there is no simple or easy answer. Consensus among observers alone will not suffice. Although often proffered as a "way out," inter-observer agreement does not verify anything but cultural coherence. That different observers agree on the placement of stress may serve to indicate only a degree of culturally mediated decision-making. Birdhwhistell's comments about the fifty million Frenchmen are apropos here, as are Mair's remarks on the validity of inter observer agreement in the study of human communicative interaction. There must be another measure; there must be an objective, physiologically based means by which to gauge what we intuit as "stress."

In his own work Scollon proposes two checks against the superimposition of expected metric pulses on whatever syllable they happen to fall. First, he marks stressed syllables independently from the audio tape and from the written transcription. Scollon argues that when the downbeats and upbeats indicated by the divide of speech into measures of duple metre coincide with syllables marked as stressed "it at least confirms that the patterns can be perceived from three separate points of view" (1981a:11). He goes on to point out that not every stressed syllable falls on the beat and that the process of "listening for metre" is not merely another way of listening for stressed syllables. Second, Scollon proposes that when the "synchronies" of an interaction "fall exactly on these downbeats and upbeats it is, I believe, convincing to argue that one has recorded the metric system being used by the original participants to organise their interaction" (1981a:11).

The present study omits the step in which stressed syllables are marked from a written transcription alone. Except for this omission, Scollon's two "checks" apply equally well to the present study and may be said to argue both for the verity and validity of the rhythmic transcriptions included in it. There are, however, major unresolved problems with the whole notion of stress and metre in speech which Scollon does not address. For example, as indicated in Chapter 2, there is as yet no satisfactory, agreed upon metric by which to measure or even mark "stress" or "pause" in speech. Researchers in this area are still trying to find acceptable ways to identify and define something we intuitively perceive and "know" is "there."

The rhythmic transcription is not essential to the argument of this work. Why, then was a decision taken to include such a potentially problematic undertaking in the study? There were several reasons, all inter-related. First, the transcription appeals to and in some sense documents the intuitively-perceived rhythm in speech. It provides a graphic, visual representation of "what everyone knows is there." Second, in appealing to and documenting the intuitively perceived rhythm underlying speech in interaction, a rhythmic transcription serves to contextualize, adding another (perhaps more concrete) dimension to the observation, the record of simultaneous manifestation of sound and movement within and between interactants. Third, a rhythmic transcription such as this may serve to corroborate the "fact" or the claim that interactants enter, exit, move, and/or speak with the pulse or beat, the rhythm, of interaction--a pulse or rhythm coherently describable by the falling S shapes of the fundamental frequency of utterance. i.e. the melody of the text. Finally, precisely because the rhythmic transcription is not essential to the argument here, problems with the basic concept or process of transcription will not and can not undermine or "taint" the present study. Advantages obtained by including a rhythmic transcription of the selected fragments of text thus far outweigh disadvantages.

Chapter 6

Selected Fragments of Text

6.1 Introduction

The notion of synchrony is central to Mair's model. It has been suggested that much of his theory in fact developed in an attempt to explain the "synchronies of all sorts, shared movements, intricate coordinations of timings and speakings" (1977:IV-15) that populate the audio and video recordings of dyadic interaction which are his data. For this study I have chosen three fragments of text for microanalysis, one from each of the three half hour conversations recorded. As a final choice it represents the culmination of a process in which criteria for the definition of synchrony were specified; Mair's "types" of synchrony were given explicit definition; recorded data were observed and analyzed over several hundred hours; in consultation with other researchers a preliminary selection of five fragments from each tape was made; and, finally, a more comprehensive analysis of these fifteen was effected in order to inform selection of a single fragment from each.

I intend to provide description of each fragment in the the terms set out in Mair's Steps Toward Principles of Text Regulation (1977). By identifying in the three selected fragments simultaneous manifestation of sound and movement within and between interactants; by documenting temporal sequencing in both audio and visual modes within and between interactants that is too swift to allow for generation by a process of feedback stimulation; by describing and analyzing these data as they occur over 0.1-second intervals; and by attending to the unique and peculiar as well as the regular and generalisable aspects of the sound and movement which constitutes the recorded data for each of the three selected fragments, I make the claim to substantiate and to provide evidence and support for the argument that synchrony is a legitimate generic term for a set of events during communicative interaction that is possessed of a specific and identifiable complex set of regularities.

Each fragment will be introduced by an overall description and then "followed along in time," since, at any one time the sound and movement clusters occur together. Interpretation, discussion of features of note from the point of view of the theory, will follow the descriptive aspect in the examination of each fragment. It is in this portion of the text where I intend to develop the specifics of each aspect of the general category "synchrony" and to outline characteristics of the set of events held to compose the general category itself.

For each fragment of the text there are the following data:

1. Graphic transcription of head and hand movement. Change of position is marked for each 0.1-second interval. For plotting head movement I chose a single spot on the face--the medial canthus of the eye: that point above

the right eye where the temple joints the "front" of eyeglasses. I followed it on successive single frames of the video tape through time. For hand movement I chose a single spot on each hand: that place where the index and middle fingers meet and join, and followed that point through time. Data points in each fragment are sequentially numbered. Numbers indicating movement positions achieved by the right hand are accompanied by the prefix "R" (e.g, "R7"); and those marking position positions of the left hand carry the prefix "L." Numbers marking head movement have no prefix (e.g., "7"). The numbers correspond to the number assigned to the tenth of a second interval in which the movement occurred.

For each fragment, 0.1-second intervals are sequentially numbered from 1 to *n*. Where there is no observable movement in a given 0.1-second interval, there is no new data point and thus no number corresponding to that 0.1-second interval in the transcription of movement for that part of the body. It should be noted that where there is a "pause" and thus "missing" numbers in the transcription, movement described by the line connecting data points occurred during the 0.1-second interval that immediately preceded the data point with the larger number. For example, if in the transcription of right hand movement there is a data point labelled "R10" followed immediately by the notation "R14," the movement which is described began at

"R13" and, moving through the 0.1-second interval achieved the position marked by "R14."

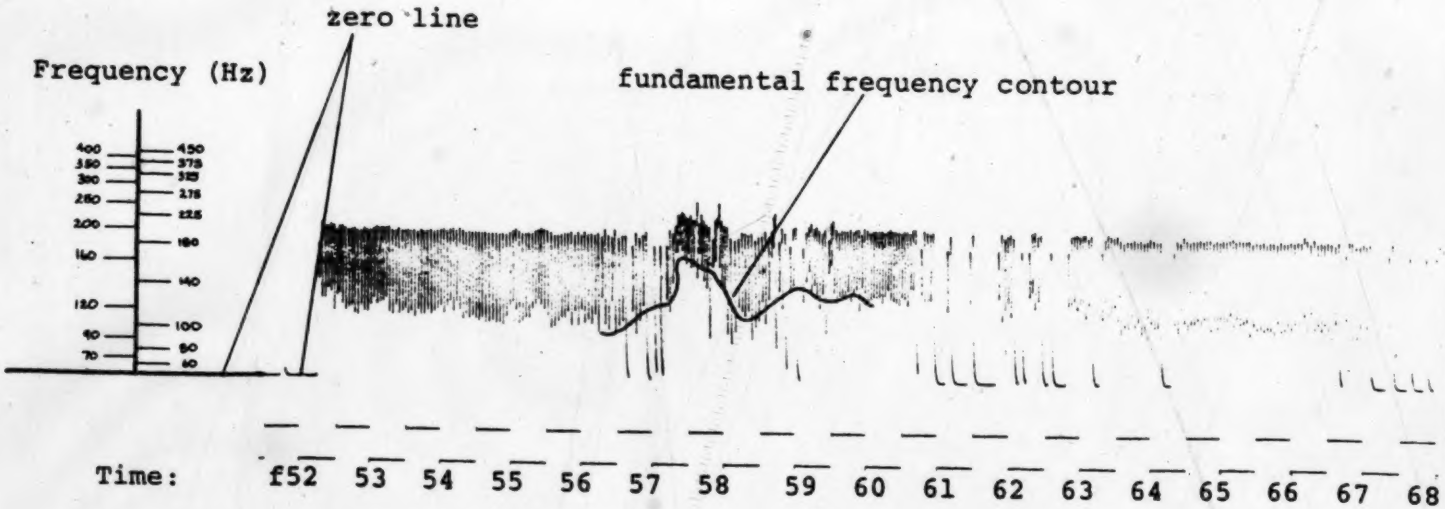
Movement is consistently described from the orientation of the subject participant. That is, "left" and "right" in transcription refer to the subject's left and right. (This makes several distinctions clear: it disambiguates left-right movement by a subject on the screen, as opposed to a "together-apart" movement as the interactants move toward and away from one another.

2. Mingograph traces of fundamental frequency of utterance. The mingograph delivers a hard copy of the fundamental frequency over time. Mingograph traces of fundamental frequency are read from the bottom up. The straight, flat line at the beginning and end of each text fragment marks what is termed the "base line." This line denotes zero frequency, silence. More specifically, when there is no voicing (i.e., when there is no vocal fold vibration during articulation of the sound), the fundamental frequency trace evidences zero amplitude. Voicing is indicated by the spikes superposed on top of the fundamental frequency curve. The vertical striations indicate larynx vibration; the fundamental frequency curve itself is the "line" traced out by the bottom of the spikes. It must be traced out manually or visually. Using the frequency calibration chart to the left of each mingograph trace, actual frequencies may be determined for any given portion of text.

The speed of the paper transport system of the mingograph is not particularly accurate or reliable or time. It is thus necessary either to construct a time scale rule and measure time for bits of each fragment, or to employ a second channel of the mingograph to record a time scale coincidentally with the fundamental frequency. The latter course was taken in this study. The time scale is evident immediately beneath the fundamental frequency trace; it is a record of the square wave generated by the pulse generator. Frequency of the square wave was precisely 10 Hz. The time scale effectively divides the fundamental frequency trace into at least potentially discrete data points, temporally equivalent to those used in preparing transcriptions of head and hand movement. The time scale is numbered identically with the time scales of head and hand movement, which themselves are derived from the time code. Integers from the mingograph time scale are accompanied by the prefix "f"--for example, "f7." Identical numbers mark identical points in time: 9, R9, L9, and f9 all mark the ninth interval of one tenth of a second in the fragment of text to which it refers. Plate 6.1 is an example of this notation.

3. Indication of voice onset and offset. Voice onset and offset are documented in the mingograph trace of fundamental frequency.
4. Transcription of utterance in standard notation.

Plate 6.1--The Mingograph Record



C: So I w a s . . . Ha!

L: Ha! They a r e n ' t s o s l o w !

"°°E.C.----- ----°°."

Transcription of utterance in standard notation is placed in time below the time scale on the mingograph record. Placement in time is accurate to within 0.1-second. It is in fact a record of voice onset and offset.

5. Graphic record of eye contact. A solid line on the mingograph record indicates the existence of shared eye contact. The commencement of an eye contact direction of regard is indicated by "°°E.C.----". Gaze aversion is indicated by "----°°". These changes in direction of regard are subjectively judged, of course.
6. Rhythmical transcription of utterance in standard Western musical notation.

6.2 Example 1

This 8.4-second fragment of text occurs approximately 24 minutes into a half-hour conversation between a female and a male, C and D. D has just introduced a new topic, that of Washington, D.C. A bit earlier there is a discussion of schools, students and teachers. Both C and D have had extensive experience in the public school system. They remark, somewhat sadly, that the school seems to have become nothing more than a place to pass the time between 8:30 and 2:30, Monday to Friday. They conclude that both students and teachers merely suffer school. Interest seems to lie in outside activities, especially full- and part-time jobs. There is some speculation as to why students in particular find it necessary to maintain the equivalent of full-time jobs, to which D responds that if you want to keep up your car payments, buy new clothes, shoes, stereos, and so forth, it is necessary to have a job. D observes that from his experiences in an inner city junior-senior high school in Toronto, the Italians are especially remarkable in this respect: "Especially the Italian guys, wanna buy custom pool cues, tight pants, new shoes with big high heels on them."

The introduction of "Italians" as topic leads to C's reminiscences of an Italian boyfriend: he also wore high-heeled shoes. More or less as an aside, C remarks that the family of her former boyfriend must have been "different" in that they were farmers in Southern Alberta. There follows a brief discussion of Italians as an ethnic group in Canada, during which C asks D, a sociologist and former resident of Toronto, if it was the case that most Italians who have come to Canada in the 1960's ended up in Toronto working largely in the construction industry. D answers in the affirmative, and, referring to their previous discussion of Toronto and his experience as a teacher in "that rotten city," launches the new topic: "Speaking of rotten cities, I was in Washington, D.C., last week." C responds with "You're kidding" and the topic is begun. C queries the reasons for and nature of the trip. There is a slight detour as they enter a brief discussion of funds available to graduate students at the university to help underwrite the expense of conferences. D then returns to the original topic and sums his perceptions of the city in the first utterance of the selected fragment, "Washington, D.C., is a sick place, man."

Following this fragment, C and D go on to a more detailed discussion of what D describes as the "sick" and "crazy" situation of Blacks and Whites in Washington. Throughout this portion of the text, D emphasizes the poverty of the Blacks and the racial discrimination they must endure. There is a long exposition on occupations and housing conditions of Blacks in the area. D observes that the effects of these circumstances mark both Black and White populations. He concludes, closing off topic: "They're sick. They're all crazy there! The Whites have nothin' to do with the Blacks, the Blacks have nothin' to do with the Whites; and they all hate each other!" C rounds off D's comments with a conclusion of her own, "Oh God, that's really depressing," and launches a new topic.

This fragment was chosen for the multitude of synchronies in evidence throughout its duration. It is a fragment populated with instances of both intra- and inter-personal synchrony. Voice and movement interweave, occurring together in smooth, closely timed sequence. D establishes the rhythm for this portion of the text with his initial utterance, and the voices and movements of both interactants follow it in time as they work at developing the new topic. During this fragment it is argued that C and D are "in rapport." They are moving through time together.

Both work at creating and organising the shared text.

6.2.1 Sound and Movement Clusters

Just prior to this fragment, D establishes eye contact with C. He has been justifying the expense of the Washington trip, describing the "nice little vacation" he got out of it. As he closes off that topic with "Who cares?", he breaks eye contact, lowering his head and looking down toward the table top. With "But anyhow," the utterance which immediately precedes this fragment, D brings his head up and toward C until he is facing her directly. He makes eye contact here, and maintains it through his first utterance, the "sick place" utterance. C breaks eye contact. She looks down as she begins the next utterance. There is no further eye contact in this fragment. During the final two or three 0.1-seconds, there is a pause in C's speech-with-movement stream, where D raises his head to look at her, but C does not re-establish eye contact.

As the fragment begins, D is sitting forward in the chair, leaning on the table. His arms lie flat on the table top, right hand on top of left. He has turned and raised his head so that he is facing C. C is sitting back in her chair, leaning on her left elbow which, in turn, is resting on the left arm of her chair. Her upper body is inclined to the left, toward D. C's right arm remains extended to the table, right hand on table top, holding her coffee cup. C is not facing D directly. Her head is turned slightly so that she is able to see and be seen by him without seeming to "confront" him head-on.

6.2.2 Text of the Selected Fragment

D:	C:
Washington D.C.	
is a sick place,	
man	I know.
	What did I read.
(whistle)	What did I just read?
That's. . .	
	It's a. . .
a sick place.	
	Inner city of Whites. . .

(Note: Utterances on the same line are produced at the same time, coincidentally. Each new line, then, indicates the next utterance in the sequence.)

Example VI.1

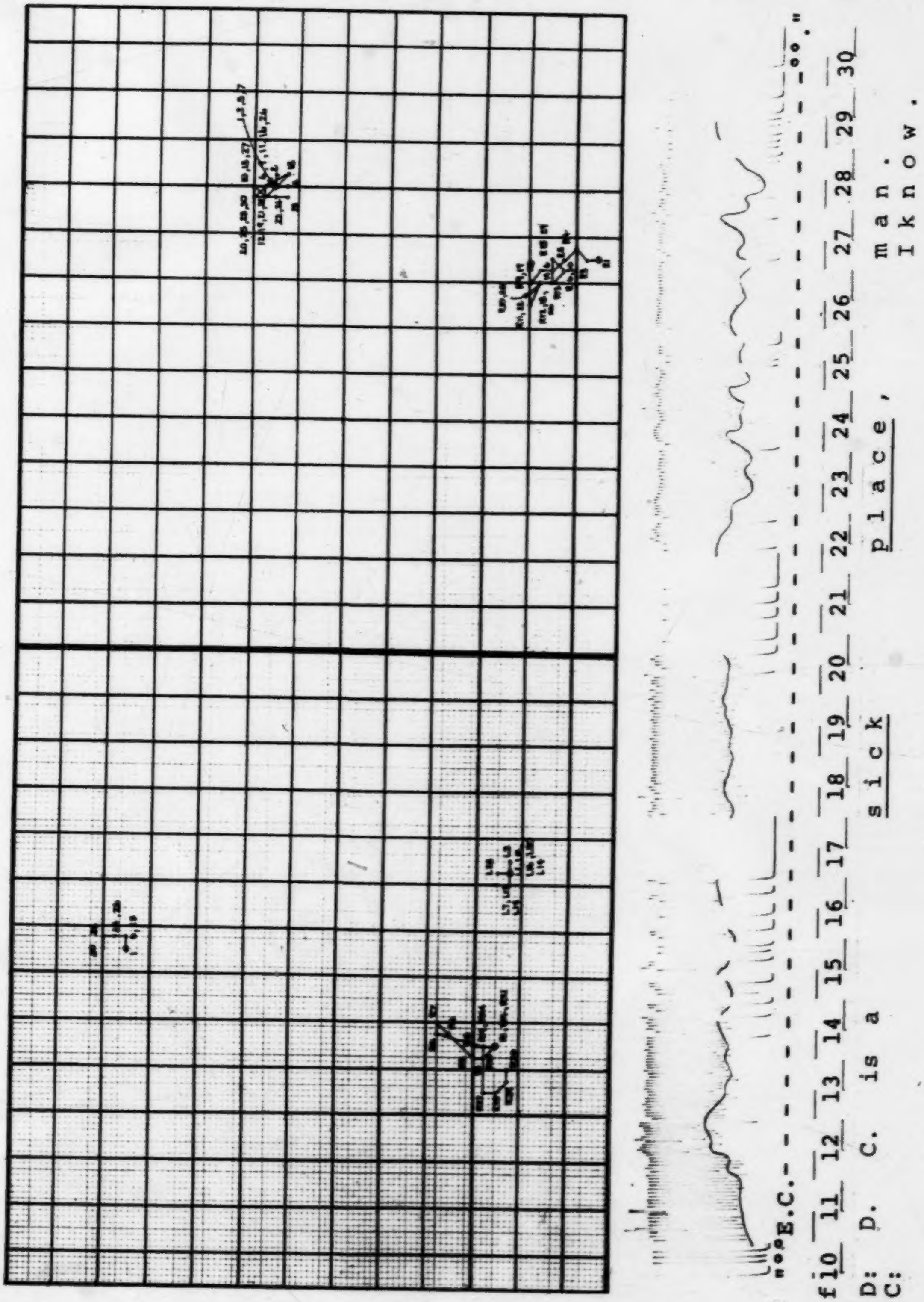
"Washington, D.C. is a sick place, man"

Example VI.1a

D's head and hand movement "go with" the movement of the fundamental frequency of utterance. Direction and extent of head movement literally follow the rising and falling S shapes of the intonation contour as revealed in the mingograph traces. There is coordination as fine as the level of the syllable. As the transcription indicates, C, in contrast, is relatively motionless during this segment. She makes only one discernable head movement, a brief lateral shift (5 to 6), and a few motions with her hands. What movement there is, though, also "goes with" the fundamental frequency of utterance.

Washington--The first syllables here coincide with lateral and vertical head movements, respectively, executed

Plate 6.2--Mingograph and Movement Traces: Example VI.1a



D: D. C. is a sick place, man.

C: I know.

by D (1 to 2; 3 to 7). The audio signal is relatively weak here; D is quite far from the microphone. As a result the mingograph trace is not very clear and it is difficult to determine the fundamental frequency for the utterance. In one instance, however, transcription of head movement and the mingograph trace indicate that the downward motion of the fundamental frequency at the end of the first syllable (f3 to 4) parallels, going with the direction of the movement of D's head at that time. D's hands remain on the table through this portion of the text. His right hand obscures much of the left, particularly that point between the index and middle fingers, used in the present study as a reference point for the transcription of hand movement. D is moving the fingers of his right hand along the top of his left throughout the segment. In colloquial terms he might be said to be stroking the back of his left hand. His right hand moves with each new syllable. Furthermore, although both head and hand may be described as moving in time with utterance, they are not exactly coincident and coetaneous movements. D's right hand, for example, continues to move through the "hesitation" in head movement noted in the transcription as D utters "...shing..." (8 to 9).

C moves coetaneously with D's "Washington." C brings her right hand up and around the handle of her coffee mug as D produces the word (R3 to R8). At the same time she shifts her left hand horizontally along the table top (L2 to L8), and makes the very small lateral head movement (5 to 6) noted above.

D.C.--These are two very distinct, clearly enunciated syllables. Subjectively, they sound slightly emphasized, but although that impression fits well with the accompanying movement there is no available measurement to corroborate it. The fundamental frequency arches slightly through "D." (f10 to 11), and rises quickly to effect a more prominent arch through "C." (f11 to 13). D's head moves down with "D." (10 to 11). Head and fundamental frequency execute parallel, coincident movements. Longer head and hand movements in this bit of the text occur coetaneously with the longer fall of the fundamental frequency to "C." Specifically, D's head goes across-right and up-left with "C." (11 to 13). His right hand continues the stroking motion, effecting what look on the video tape to be more pronounced movements as the fundamental frequency falls.

C also moves with the fundamental frequency of utterance here. On "C.", right hand movement imitates the direction of motion of the fundamental, going with it. C brings her right hand down to rest on the table, retaining her hold on the handle of her coffee cup (R12 to R14). Her left hand makes a series of right-left-right movements through this, changing direction with the fundamental frequency of "C."

There is a similar coincidence of sound and motion during "is a." C's left hand changes direction with the production of each word syllable (L12 to L16).

Sick Place--are long, drawn-out and emphatic word syllables. D's head movement is strong, and definite. Head movement seems emphatically to punctuate the utterance with a kind of kinesic grammar. D brings his head down on each of the stressed syllables. Head movement is more pronounced on the latter syllable, "place," and exactly parallels the down-up motion of the fundamental frequency contour (f21 to 25). Longer movements occur with the emphatic pulse of each syllable. Finally, it is perhaps important to note that there are three 0.1-second pauses or hesitations in the movement of D's hand during "sick place." Movement following each pause tends to be longer, as if to arrive at the same point in time as the hand(s) would have got to had there been no interruption in the flow of the motion.

C continues to effect small hand movements in time with fundamental frequency. With "sick," C moves her right hand up the handle of her cup (R18 to R19). She initiates the upward motion of her right hand precisely as the fundamental frequency peaks and reverses direction. "Place" is marked by two motions of her right hand: as the utterance begins (R20 to R21), C moves her right hand down along the handle of her cup; and, as D brings the fundamental frequency down, she brings her hand up the cup handle once again (R23 to R24). Right hand movement changes direction with the fundamental frequency.

The final word/syllable of D's utterance, "man," is barely audible, largely obscured on the audio-tape by C's quick "I know." It seems safe to assume that the mingograph trace here (f26 to 30) marks C's utterance: hers is the louder and stronger, and is more clearly reproduced by the audio recording equipment. From the mingograph trace it is apparent that C enters the discussion on the last beat, i.e., in time with the rhythm established by D's utterance. This and subsequent utterances demonstrate that she has taken on the established rhythm of the text. Of especial import for the present study is the fact that C's utterance occurs with D's, entering within 0.1-second of the initiation of his "man."

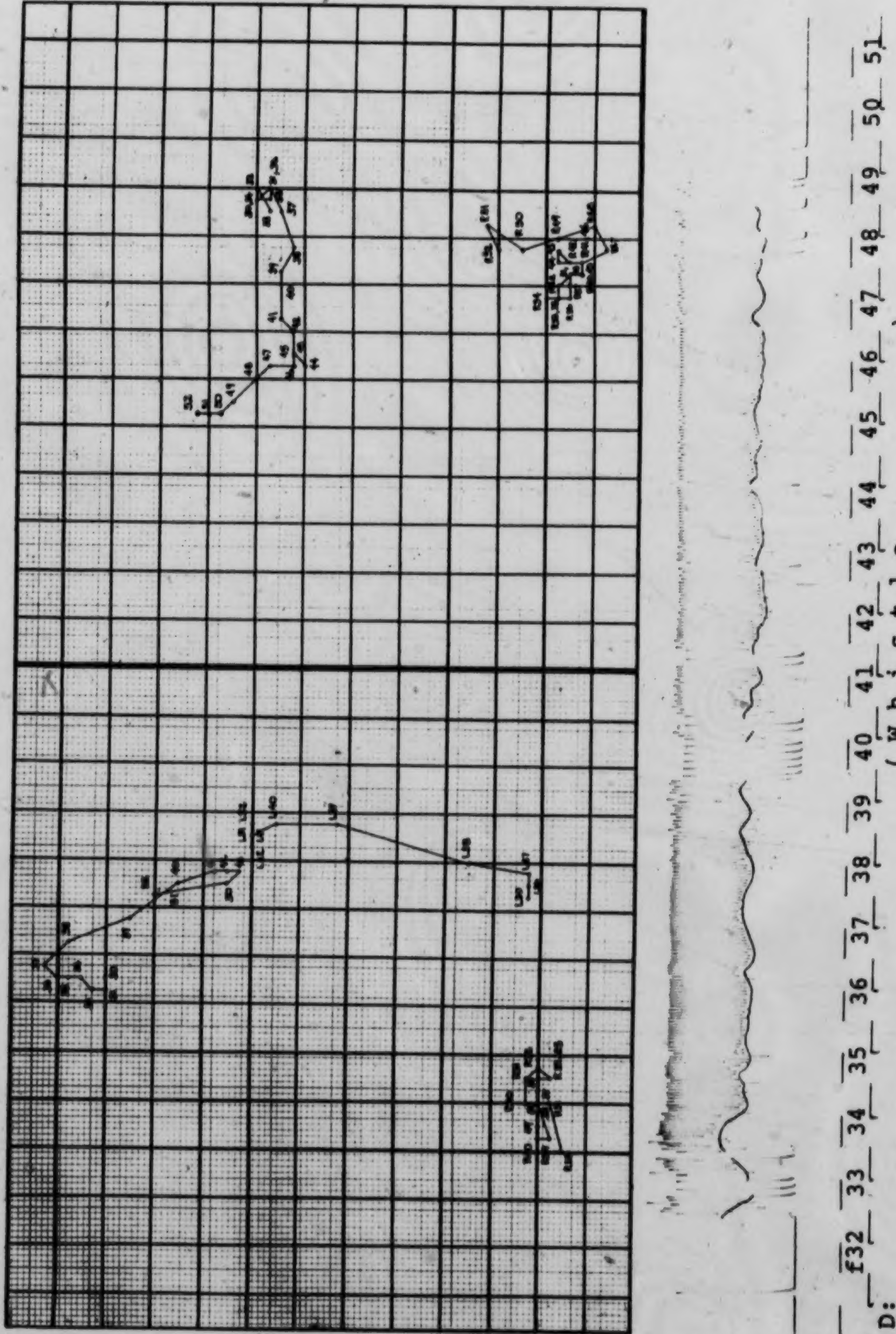
```
I know. What did I read. What did I just read?
...man.                (Whistle------------)
```

Example VI.1b

In the instant between D's and her own utterance, C breaks eye contact, looking downward to the table top and her hands. Although there is no further eye contact, the segment is marked by striking examples of coincidental and coetaneous clusters of sound and motion. C and D move and speak in time with one another, with the rhythm of the fundamental frequency of utterance.

I know--C begins the head movement that goes with this utterance almost half a second before the utterance itself is produced. She begins a nodding motion as D produces "place," and continues the motion to within 0.2-second of the completion of her utterance (28). Coinstantaneously with D's completion of "place," C effects a horizontal left to

Plate 6.3--Mingograph and Movement Traces: Example VI.1b



right movement with her right hand, shifting away from the handle of her cup (R25 to R27). This hand movement turns down-right precisely as she rounds off "I know" (R27 to R30). At the same time, she raises and lowers her left hand (L28 to L30). These left hand movements parallel the rise and fall of the intonation contour (f27 to 30).

There is a moment of silence between utterances. In that moment C makes a horizontal, left to right head shift (29 to 32), and coincidental, coetaneous down-right, left-across right hand motion (R30 to R32).

What did I just read--is marked by a number of simultaneous events in both aural and visual modes. "What" has an arching fundamental frequency (f32 to 34) that takes up the trajectory where it would have got to had there been no silence, and loops it over the top. "Did I" has a relatively "flat" intonation contour but continues the downward movement (f34 to 37). The mingograph trace indicates a "hiccup" in the fundamental frequency contour where "I" is heard on the tape (f36-37). On "read" the fundamental frequency describes a complete arch through the first part of this word-syllable, then turns to move upward again into the silence marking completion of the utterance (f38 to 40). The entire sentence is accompanied by a congregation of movements. Specifically, C looks up toward the ceiling (32); raises her head, tilting it back (32 to 37); plants her right hand on the table, fingers spread and arched (R32 to R36); and shifts her left hand across-left

(L30 to L37). This complex is accomplished in one swift, continuous confluence of sound and motion. The eyes flick up; the head lifts and tilts; the utterance is voiced; the right hand comes down; the left hand shifts. It all happens as one, at once.

D is part of this as well. He moves in time with C. As she lifts her head, leaning it back, D's head describes a pattern that may be likened to an "X" in a three-sided box, open on the left (32 to 37). He begins at the bottom left corner, moving up-right along the diagonal and left across the top as the fundamental frequency arches through "what"; slips down-right along the second diagonal and proceeds up the right side, sliding back down the diagonal to the point of origin with "did I." At the same time, D's right hand moves quickly across to the right and up, following the rise and fall of the fundamental frequency.

It is with "read" (f37 to 40) that C and D begin quite obviously to move together in time. As C moves along the vertical axis, D moves along the horizontal. C and D. coinstantaneously initiate sequences of movement that go with the fundamental frequency of utterance and the movement of the other. As the utterance begins (37), C is at the apex of the arc described by head movement. Through the utterance C moves her head down-left. The movement continues past the end of "read" into the next segment, where she places her left hand over her eyes (41-42). In the same motion, at the same time, C moves both hands: she raises her left hand from the table top toward her eyes (L37 to L41); and she shifts her right hand to the right, collapsing the hand palm down on the table top where it remains, motionless, for almost a full second. D meanwhile initiates a left to right lateral head swing (37 to 46). He lowers his head, tilting it forward, and, looking down toward the table, swings his head to the right. The movement of D's head describes a fluid, pendulum-like arc. Movement proceeds through "read" and on into the succeeding segment. It must be emphasized that these are all one motion; on the video screen they appear as a single continuous flow of sound and motion. It all happens together, taking no more than three or four tenths of a second. Such fluid coincidence is difficult to reproduce in the description.

C produces most of "what did I just read" from "behind" her left hand. She achieves her behind-the-hand position with "what" (41), and remains there, motionless, until the last word in the segment. The fundamental frequency of C's utterance and that of D's whistle are almost inextricably intertwined on the mingograph trace. It is possible, however, to disentangle the two sufficiently to provide a highly probable description of the fundamental frequency of each vocal production; that is, one that agrees with and appeals to what is heard on the audio tape. On "what" the fundamental frequency picks up from where it would have got to had there been no silence, and rises slightly through most of the syllable, falling just before the "t" (f40 to

41-42). The balance of the utterance has a fairly flat contour: there is a slight rise with "did I" (f42 to 44), and similar slight drop through "just" (f44 to 46-47). With "read" the fundamental frequency is turned around carried up into the higher frequencies that here mark "question" (f46-47 to 50).

D's whistle begins one tenth of a second before C's utterance. It appears to the ear the fundamental frequency of the whistle describes a long rising-falling arc. It is the mirror image of the arc described by the movement of D's head over the same time period. D's right hand moves with, "embroidering" with smaller movements the large head movement. (Right hand movement is down-left [R37 to R39], across-right [R39 to R40], back left and up [R40 to R43], and down again [R44 to R46], ending a little down and to the left of where movement began.) The stroking motion of his right hand ends here. D begins to ready for a major body shift (described in detail below.) As noted above, D looks down just as he begins to move his head through the lateral arc, and does not bring his gaze back up until very near the end of the fragment (80) when, in the silence of C's pause, he swings his head up and brings his gaze to bear directly on her once again. (Exact direction of regard during this portion of the text is not noted here. The relative positioning of head and camera make it impossible to observe eyes and eye movement.)

It is interesting to note that the longer, larger movements of both interactants occur through the silence between C's utterances (38 to 41).

As D's whistle ends, he begins to raise his head (46 to 49), and moves his right hand forward and on to the table top (R46 to R49). These mark the preliminary stages of the major body shift that follows, and coincide with, in fact paralleling, the motion of the fundamental frequency of utterance produced here. As C utters "read" (f47 to 50), she begins to move out from behind her left hand (49 to 50). At this point C also commences a clockwise circular motion on the table top with her right hand (R49). She continues the movement, circling without stopping, for almost 2.5 seconds (R49 to R72), until she produces "inner city" and completes the state of play of which this movement seems to be an integral part.

D's body shift carries on into the next set of utterances, extending through to the completion of "that's" (46 to 57). The body shift is indicated in the longer movement of hand and body on the transcription. D straightens in his chair, sitting up to the right, and then comes back down towards the table and a position of "repose," arms resting on the table top. Head movement describes a large loop which reaches its apex in the centre of the silence between utterances (52 to 54). This movement occurs in time with the rhythmic pulse created by the melody of the text. D lifts his hands and shifts them along the table top toward his body during this same time period (R49 to R57). Movement and utterance are completed within 0.1-second of one another. Finally, as C rounds up "read," D lifts his head up and to the right, looking away from C (48 to 53). On the video tape he appears to flick his head up precisely as C moves out from behind her left hand and says "read."

Again, it is noteworthy that the longer, larger movements occur across and through the silence between utterances.

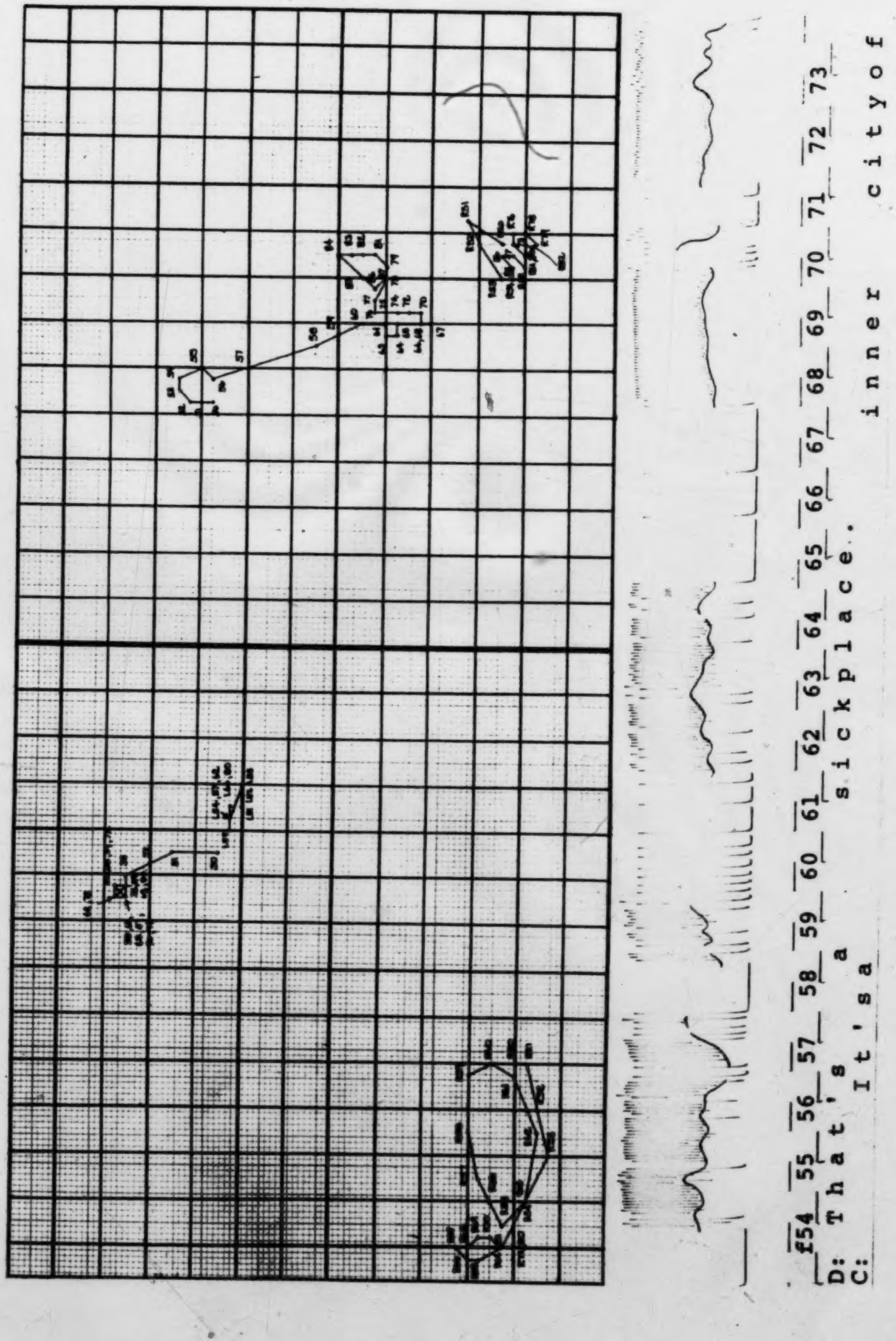
```
That's . . . a sick place.
   It's a  . . . .         inner city of Whites
                                        Example VI.1c
```

This segment overlaps with and carries on from the previous. As well as demonstrating the arbitrary nature of the divide of the speech with movement stream into segments for the purpose of analysis, this segment provides further instances of complex coincident and coetaneous production of sound and movement within and between interactants.

An important aspect of this segment occurs as it is begun: both interactants come in together after the silence between utterances. C's "it's a" comes in 0.1-second after D's "that's." C's utterance comes in literally on top of D's. On the tape, listening to it in real time, she drowns him out almost completely.

<u>That's</u>--As already noted, D is in the process of effecting a body shift as this series of utterances begins. He produces the softly spoken "that's" (f54 to 55) as he

Plate 6.4--Mingograph and Movement Traces: Example VI.1c



brings his head up to the right, across-left, and down (51 to 55). This is the "flick" described above. The arc described by his head is almost identical in shape and duration to that of the fundamental frequency contour. Again it must be noted that a two-dimensional transcription of a single point on the face is patently inadequate to the reality. In this instance the plot cannot properly indicate inclination or rotation of the head; nor can it mark movement "into" or "out of" the screen.

It's a--the fundamental frequency falls through "it's" (f55 to 56-57) and quickly rises again through "a" (f56-57 to 58). C is, also, completing a sequence of movements as this segment begins. During this utterance C comes out from behind her left hand, completing the movement, and rests left cheek on left hand (55 to 57). Once C achieves this position, left hand and head remain quite still for the balance of the segment. Any further movement is very small, almost unplottable. The plottable bits do, however, continue to move in time with the fundamental frequency of utterance. As the fundamental frequency falls on "place" (f63), for example, C moves down-left with it.

C's right hand continues the circular motion on the table top. And, once again, longer movement occurs through the silence between utterances (61 to 62).

A sick place--the fundamental frequency rises slightly through "a" (f59), and continues across the "silence" marked on the mingograph trace to be taken up by "sick" (f61-62).

(The "silence" is perhaps the voiceless /s/ which, having no fundamental frequency, cannot be properly recorded with this trans pitchmeter-mingograph set-up.) The fundamental frequency continues to rise through "sick" and falls, tumbling over in a double arch, through "place" (f62 to f66). While he produces this phrase, D inclines his head forward and brings it down (moving downward from the apex of the arc of the "flick"). In the end he is leaning forward, looking directly down toward the table top. These movements parallel the motion of the fundamental frequency. They are coincidental and congruous with it. Furthermore, the motion in both aural and visual modes terminates in the same 0.1-second, with a pause. It is notable that D's hands remain virtually motionless during this portion of the text. They rest on the table and effect, as the transcription indicates (R58 to R66), only very small movements not immediately relatable to the movement of the fundamental frequency or other body movements.

C, too, is quite still during this bit. The paucity of her movement has already been remarked upon, as have the pertinent movements of the constellation of head and left hand. Her right hand continues the clockwise circular motion.

Inner city of Whites--C evidences a marked lack of physical movement during this utterance. It is atypical of her interaction style. There are small movements of the head/left hand combination, as indicated by the box-like formation of lines on the transcription (68 to 80), and two other movements that are of note. But that is all. D is more active. His hands remain relatively still (R66 to end), but there is no lack of head and body motion.

Just prior to the initiation of "inner," C stops the clockwise circular motion of her right hand (R65), and brings her hand down to rest on the table top. A tenth of a second later, but still in the silence between utterances, C raises her right index finger and taps it once gently on the table (R66 to R68).

Inner--the fundamental frequency is taken up and continues to rise through the first syllable, but falls through the second (f67-68 to 71). With "city" the trajectory of the intonation contour is picked up across the silence. Again, the fundamental frequency rises through the first syllable of the utterance, tumbling over and down with the second, completing the wave form (f72 to 74). Generally, the fundamental frequency may be said to rise in a double arch through "inner city," falling off only with the last syllable produced. As C utters "inner city," she brings her right index finger back into her loosely closed fist (R69 to R72). Particularly when viewed on the video tape, in real time scale, this appears to be a motion sequence that goes with utterance, punctuating sound with movement. For the balance of the fragment, C's right hand remains motionless on the table top. C's head/left hand constellation executes several "small" movements through this portion of the text, but they are too fine to be meaningfully transcribed here.

D moves with C's "inner city," bringing his head up and to the left in time with the generally rising fundamental frequency of the utterance (66 to 75). He maintains a downward gaze throughout.

Of whites--the fundamental frequency arcs up through the utterance, reaching the apogee with the long vowel of "whites" (f77), and dropping off rapidly at the end. Silence fills the final moments of the fragment (f80 to 87). C brings her gaze up as she produces "of." Her eyes flick toward the ceiling. Corresponding, coinstantaneous, head movement is a slight shift up-left and then right (71 to 73). From this point in time to the end of the utterance and beyond the limits of this fragment of text, C is motionless. She appears on the video tape to be staring, fixating on a point in space at eye level and slightly to her right. C's last movement of note is not recorded in the transcription, although it is obvious on the screen. It occurs with "whites" and consists of a side to side (left, right, left, right) flicker of her eyes. The movement seems to be part of, going with and adding dimension to, a later state of play--that delivered with the utterance "surrounded by Blacks and they're scared shitless," which immediately follows the fragment selected here.

As C begins production of "of Whites," D begins a second body shift, but on a smaller scale than the previous. He moves his torso and shoulders up, forward, and over the table. As evidenced in the the transcription, his head movement, up, across-left, and down-left (74 to 79), parallels the arc described by the fundamental frequency. Subsequent head movement consists of a loop up and around to the left (80 to 87). D brings his gaze to bear directly on C as he completes the loop, but C is still "staring out into space" and, as noted above, does not re-establish eye contact.

6.2.3 Sound and Movement Clusters: Interpretation

This is a very rich fragment for the present investigation of synchrony. Throughout the fragment there are in evidence synchronies of all sorts: shared movements and timings, complex interweaving of melody, completion of utterance of one by the other, coincident and often coetaneous production of sound and motion both within and between interactants. Such intricate interlacing of speech and movement argues strongly for the notion of synchrony. As with all the examples, the above description, although detailed and lengthy, hardly begins to convey adequately the complexity of events as they happened. Transcription of a single point on the face and one spot on each hand are not up to the detailed interplay of sound and motion within and between interactants that constituted the interaction "as it happened."

In the following sections I interpret the events described from the point of view of the theory. I shall argue that the constellations of sound and movement identified constitute evidence for the existence of inter- and intrapersonal synchrony in human communicative interaction, that intra- and interpersonal synchrony are in fact identifiable, definable, observable aspects of the communicative process. For each segment, then, the focus will be on those sound and movement clusters thought to provide evidence for the existence of synchrony in human interaction.

"Washington, D.C. is a sick place, man"

Example VI.1a

Washington--The fine integration of head and hand movement with fundamental frequency contour is taken as evidence for the validity of the notion of interpersonal synchrony. Such coordination between modalities is, in addition, thought to be the product of a single central patterning program in the brain. Movement and sound occur together in time; they begin and end coincidentally, within 0.1-second of one another. The very small time frame does not allow for a feedback loop between visual and aural modes such as might be argued to account for the observed synergy. Nor can movement be considered an addendum, intended merely to embellish the semantic point of the lexical utterance. It all happens too quickly. Sound and motion are too finely integrated. They occur coincidentally and coetaneously, moving through states of instability to stability in time. Specifically, D moves his head down with the fundamental frequency in "Wa...". His right hand moves in time with the pulse of "Washing...". The analogy of a differential gear might be invoked. D's head and hand move in time with the utterance, but they do not always perform parallel, identical movements. For example, there is a momentary pause in the movement of D's head at "...shing..." that is not mirrored in either right hand movement or fundamental frequency of utterance. Sound and motion seem to be more adequately accounted for as variable aspects of the same thing, products of a single central patterning program in the brain.

D.C.--The concise, angular movement of D's head goes with the clear, precise (almost emphatic) enunciation of these word/syllables. Head movement parallels the motion of the fundamental frequency of "D." Longer, larger movements of head and hand occur with the falling fundamental frequency through "C." Both aural and visual modes thus move together to a point of stability in time. These data are also thought to provide support for the notion of intrapersonal synchrony.

Sick place, man--D moves his head with the fundamental frequency as he produces this bit of utterance, using physical movement to emphasize and mark what is already, at the same time, marked by the falling fundamental frequency, tone and amplitude of voice. More pronounced movements of head and hand go with the falling fundamental frequency of each syllable/word. This simultaneous congruence of sound and motion within an interactant is argued to be evidence of intrapersonal synchrony in action. All modalities of the system, D, appear to be moving together in time and through time. In addition, the longer hand movements that seem to move to "catch up" what has been lost in momentary pauses would seem to add support for the concomitant argument for a single central patterning program in the brain as generator (or regulator) or the speech with movement stream.

The argument for interpersonal synchrony, and coincidentally, for the notion of interaction as a single system in evolution of which interactants are nodes, organising and organised centres, is thought to receive support from the fact of C's movement in time with D's utterance to such an extent and fine degree of coordination that they can only be described as "synchronous." It is thought that she is immersed in the evolving model to the point that she is taking the timing of her own cognitive processes, including actions, from rhythms impinging on her senses, D's utterance. Her movements imitate the direction of motion of the intonation contour. She moves with utterance at moments of emphasis, whether that emphasis is indicated by volume or falling fundamental frequency. And she follows along in time with the verbal text to point(s) of momentary stability in time. Of note are the movements of her right hand along the handle of her coffee cup. This series of movements parallels the motion of the fundamental frequency of D's "Washing...", "D.C.," and "sick place," rising, falling, and pausing with it.

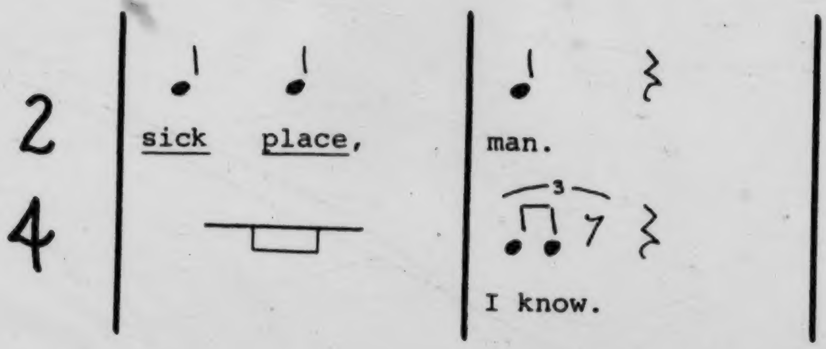
C's "I know" comes at precisely the same time as, in perfect rhythm and harmony with, and literally on top of D's "man." On the mingograph trace her utterance is virtually indistinguishable from D's, even though the fundamental frequency of her voice is much higher than his. She matches the established rhythm and melody perfectly. Such simultaneous production of sound between interactants is argued to strongly support the notion of interpersonal synchrony. Synchrony seems to account most accurately and adequately for the coincidences observed. Both C and D appear to be at some level aware of the fundamental frequency of voice and the rhythm it establishes (or is evidence of) in utterance, in interaction, and they use that knowledge. In this case they enter the conversation at precisely the same moment in time and in the same place of the trajectory of the intonational contour. Figure 6.1 is a representation of this portion of the text in standard Western music notation.

```
I know. What did I read. What did I just read?
...man.                 (Whistle------------)
```

Example VI.1b

I know--There are numerous instances of what is interpreted as intrapersonal synchrony during this utterance as well. There is close coordination of movement and fundamental frequency within C to a degree that obviates claims of a stimulus-response model as generating source. Of

Figure 6.1--Musical Notation of "Sick place, man"



particular note are the downward movements of both right and left hand that go and stop with the falling intonation contour of "know."

What did I read--The simultaneous occurrence of C's utterance with a complex of movement is thought to constitute striking evidence of intrapersonal synchrony. It is a remarkable sequence. Everything happens together. The intonation contour describes a complete sine wave figure; the right hand comes down and is planted on the table; the head is raised and tilted back; the eyes flick up; the left hand shifts. It is all over in less than half a second. There is not time for a feedback mechanism to operate. C does it all at once. It is the coincidental and coetaneous nature of this manifestation of sound and movement that is argued to be intra personal synchrony. Such coincidence across modalities within one interactant is also thought to be the manifestation of a plan that is parceled out through the various sensory modalities via a single central patterning program.

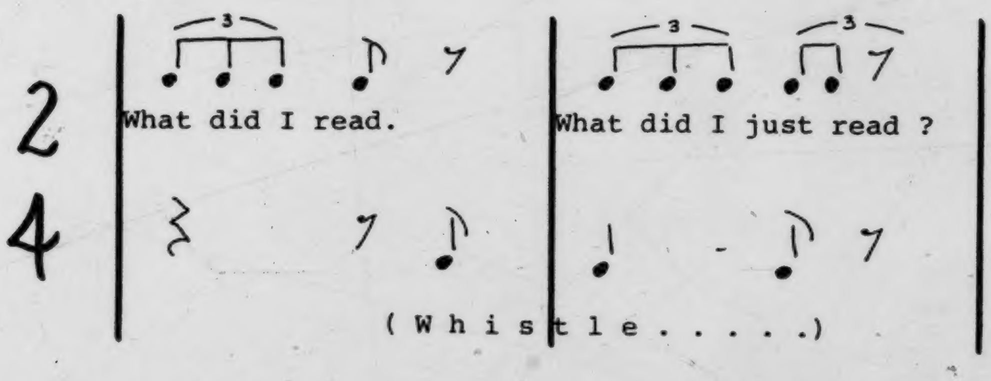
D takes full part in this synchronous activity. D moves with C. As C lifts her head, tilting it back, D moves his head through a series of box-like movements. D also moves with the fundamental frequency of utterance. In particular, D's head and hand parallel the rise and fall of the fundamental frequency through "What did I." The simultaneous manifestation of sound and movement between interactants is taken as evidence of interpersonal synchrony. To repeat, it is argued that D is in fact immersed in the shared text, taking the timing of his own cognitive processes from the rhythms impinging on his senses; that is, C's utterances.

It is with "read" that C and D begin very obviously to move together, coincidentally and coetaneously, in time. He moves along the horizontal axis precisely as C moves along the vertical axis. As C shifts her right hand and brings her head and left hand toward one another (in the end placing her left hand over her eyes), D begins the lateral swing of his head. (This head movement continues through the silence between utterances, on into the succeeding segment.) C and D move together through this. Furthermore, the movement of both goes with the fundamental frequency of utterance, following not only the direction of motion, but the pulse or beat of utterance. C and D seem to "take off" with the falling fundamental frequency in "read" and follow it on up. It is a "symphony" of sound and motion, one that is thought to provide a formidable array of evidence for the notion of interpersonal synchrony. (The events in and around this utterance stream are in fact the reason that this fragment was initially chosen for micro analysis.)

C and D continue their "dance" through the silence between utterances, going with the implied trajectory of the intonation contour across the silence. Their movement is described below. In terms of the intonation contour, D catches with his whistle the rising fundamental frequency from "what did I read" at the point to where it would have got had there been no silence, and begins to bring it down through the falling S shape at the same time as C begins her next utterance. It, too, turns the intonation contour over and down. The whole is a joint production, and is argued to strongly support the notion of interpersonal synchrony in interaction. Figure 6.2 is an attempt to represent this in standard Western music notation. It is possible to present the shared rhythm of this "dance" of conversation, the rhythm created or evidenced by the fundamental frequency of utterance.

What did I just read--This segment carries on from the previous, overlapping with it. There is manifest evidence of synchronies of all kinds here. Most apparent are the instances of what are here interpreted as interpersonal synchrony. C and D come in together after silence and continue to move and speak together in time through the segment. D's whistle comes in 0.1 second before C's utterance. C brings her head down to her left hand, covers

Figure 6.2--Musical Notation of "What did I read"



her eyes, and produces "What did I just read" at the same time as D looks down, whistles, and completes the lateral swing of his head. It all happens in one fluid rush of sound and motion, persisting over almost a full second. It is this constellation of coincident coetaneous sound and movement that is here argued to be interpersonal synchrony in interaction. The temporal sequencing between interactants is too fine and too detailed to be the product of a stimulus-response type or feedback model. There is no time. The sound and movement are generated by both interactants, together in time. On the video screen C and D appear as partners in an intricate, precisely-timed and well-rehearsed dance. But this dance is spontaneous; they create it as they go, in time with a melody, and rhythm, that they themselves must create (and in this case it is accomplished without eye contact). The aural and visual modes of both work to create the shared text. They move and speak coinstantaneously, beginning, ending, shifting movement and utterance within

0.1 second of each other.

Small movements of both interactants also seem to be co-patterned, made manifest in time with and as part of the melody of the text. As C produces "read" and pulls out from behind her left hand, D lifts his head up slightly and to the right. It is here that he makes what I have termed the flick of his head.[12] C and D move with each other here, going with and paralleling the motion of the fundamental frequency. These coincidences of sound and movement are interpreted as evidence of interpersonal synchrony. C and D are seen as proceeding together in and through time.

(Whistle...)--D's whistle describes an arc in the mingograph trace of fundamental frequency that is the mirror image of the arc described by his head through the same time period. Right hand movement also goes with, generally paralleling, the direction of the motion of the fundamental frequency. Movement also seems to go with the rhythm of utterance. That rhythm is represented in the rhythmic transcription of this portion of the text in Figure 6.2. In addition, as he begins the lateral movement of his head, D shifts his gaze downward. This simultaneous manifestation of sound and movement within an interactant is argued to be an instance of intrapersonal synchrony in action.

What did I just read--C achieves her behind-the-hand position with "what." She moves out and up from behind her hand with "read," the movement going with the rising fundamental frequency. At the same time, C begins the clockwise circular motion with her right hand on the table top. These coincidences of sound and movement are here interpreted as evidence of intrapersonal synchrony.

---

[12] It must be remembered that due to the mass of the object involved, this movement is must slower and more ponderous than a comparable movement of the hand or eye.

```
That's . . . a sick place.
   It's a  . . . .          inner city of Whites
```

Example VI.1c

That's--With the completion of this word/syllable D accomplishes the large body shift undertaken with the last word of the previous segment (i.e., "read"). The coincident coetaneous nature of the arc described by the movement of D's head and the fundamental frequency through time argues strongly for the concept of intrapersonal synchrony. It happens far too quickly and with far too much precision to be the product of a feedback loop between modalities, or worse yet, an accident.

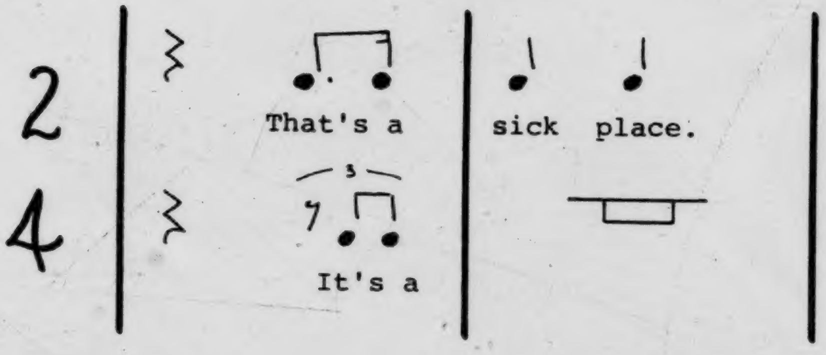
It's a--C moves with the fundamental frequency of her utterance, completing the up-right motion that carries her out from behind her left hand as the fundamental frequency rises through "it's," and settling, left cheek on left hand, as the fundamental frequency slips down through "a." It is argued that these coincidences of sound and movement constitute evidence for the existence of intrapersonal synchrony in interaction.

The simultaneous coming in together after silence in evidence as this segment begins is thought to be an example of interpersonal synchrony. C and D come in together, within

0.1 second of one another, following the pause of silence that succeeds's C's "what did I just read." Both interactants seem to be aware of an underlying rhythm or beat, and they enter simultaneously on that beat. Listening to the tape (at real time speed) it sounds as though both utterances come at the same time, C's "It's a" drowning out D's "that's." The ensuing utterance, D's "a sick place," carries on from C's interruption without missing a stroke. D catches the trajectory of the fundamental frequency contour and picks up the beat, exactly as it should be, as it would have been without the intrusion of C's utterance. It is proposed that this is in large part due to the fact that C's utterance itself occurs on the beat, and follows the projected trajectory of the fundamental frequency so completely. Using standard Western musical notation, it might be possible to represent the beat, or rhythm, of the intonation contour through this portion of the text, and Figure 6.3 is an attempt to do so.

A sick place--D rounds off this utterance with a descending S-shaped intonation contour. The intonation curve is produced coincidentally and coetaneously with a small body shift that consists of a forward and down movement of head and shoulders. In the silence that marks the completion of his utterance, he pauses for roughly 0.1 second, ceasing all motion. These simultaneous occurrences are here interpreted as manifest examples of intrapersonal synchrony. Once again, the time frame and precision of movement and

Figure 6.3--Musical Notation of "That's a sick place"



utterance production argue for the adequacy of the notion of synchrony.

Inter personal synchrony through this portion of the text is thought to be evident in the congruence of motions effected by the fundamental frequency, and C. As the fundamental frequency falls through "place," C's left hand configuration moves down briefly, in time with the fundamental frequency.

Inner city--The fact that this utterance is produced as C stops the circular motion of her right hand and gently taps the table top with her right index finger is thought to support the notion of intrapersonal synchrony in interaction. These movements occur with, and as integral parts of, the state of play being delivered with their occurrence. It is further proposed that they are most adequately and accurately portrayed as synchronous manifestations of a single central patterning program in the brain.

D's head movement here goes with the fundamental frequency of C's utterance, paralleling its movement. Such coincidence, such congruence of sound and motion between interactants is here interpreted as interpersonal synchrony.

Of Whites--C looks up quickly, eyes darting toward the ceiling, head shifting, as she begins this utterance. The timing and direction of her movements match those of the fundamental frequency. D also moves with the fundamental frequency. He effects a small body shift that parallels the rise and fall of the fundamental frequency contour. The former coincidence of sound and motion is interpreted as evidence of intrapersonal synchrony; the latter of interpersonal synchrony.

6.3 Example 2

This 11.4-second fragment of text occurs roughly twenty minutes into a conversation between two males, D and E. They are talking about a questionnaire which D had wanted to include in his Ph.D. dissertation project. It is clear that D wanted the questionnaire because he felt it was necessary to the satisfy the demands of "academia," not because it was imperative to his research. The topic begins after a lengthy discussion about the thesis versus the non-thesis route to a masters degree in their department. D and E talk at some length about a particular project that had been recently submitted in partial fulfillment of the requirements of a non-thesis M.Ed. The project was apparently excellent, considered "better than some Ph.D. theses we've had in this department." Mention of the department's evaluation of the project leads to discussion of the departmental evaluation of E's own M.Ed. thesis and the question "What is a thesis supposed to be?". D responds with a brief statement of the thesis requirements of the group with which he works in the department.

This discussion of "theses" leads directly into the topic of the questionnaire. D leads off: "Did I complain to you already about, uh, trying to do my questionnaire?". D "complains" that his committee has vetoed the idea of the questionnaire. They will allow him to do it if he wants to, but, according to D, they insist that the questionnaire is basically "dumb." They argue that D will get "much richer" data from interviews. D expresses surprise that the committee has taken such a strong stand. They do not want the questionnaire even as an appendix, for statistical corroboration of other findings. D is not sure that completely excluding the questionnaire is such a wise move. He maintains that "traditional standards" and expectations in the sociology of education, and the common knowledge that "it's easier to get a job if you can do that statistics stuff" would seem to argue for the inclusion of the questionnaire somewhere in his dissertation.

E shares D's surprise at the committee's position and wonders at the wisdom of their decision. He is prepared to admit that the questionnaire is "dumb" and of no real value to the argument of the dissertation, but he argues that "you need it" or "you leave yourself open to attack" on the grounds of lack of any "real strong evidence" from the "empirical point of view." As the fragment begins, D is in the process of adding a new dimension to the argument for the questionnaire. He contends that "that questionnaire--could get me a. . ." and the fragment is self-explanatory from there.

The text following this fragment carries on the discussion of the merits of D's questionnaire. E is not convinced that D ought to omit the questionnaire altogether. He argues that, if nothing else, the questionnaire will serve to demonstrate the inadequacy of the questionnaire/statistics methodology for this type of research. Such a finding, he suggests, is important in and of itself, and has serious implications for much of the educational research at present underway in the Department of Education. D concurs, but then admits that his committee has more or less "convinced" him. And, because the questionnaire would be more work right now, he is prepared to restrict himself to the interview methodology and leave the questionnaire and statistical manipulation as something he can do in the future. He is resigned if not satisfied.

E is not convinced, but seems willing to let the topic go. He rounds it off with a long and drawn out "I don't know," and launches the new topic: "Well, you're gonna be finished by the end of this year, eh?".

The following fragment evidences numerous instances of both inter- and intrapersonal synchrony. In particular the four to five seconds in the middle of the fragment manifest a complex interweaving of sound and motion whose intricacy almost defies description: the "dance" must be seen and heard to be really appreciated. As in the first example, it is argued that the interactants are "in rapport" here. They are thought to be immersed in the evolving text, at work developing topic together. D establishes the basic rhythm for this fragment of text with his initial string of utterances. And the movement and utterance of both follow it in time.

6.3.1 Text of the Selected Fragment

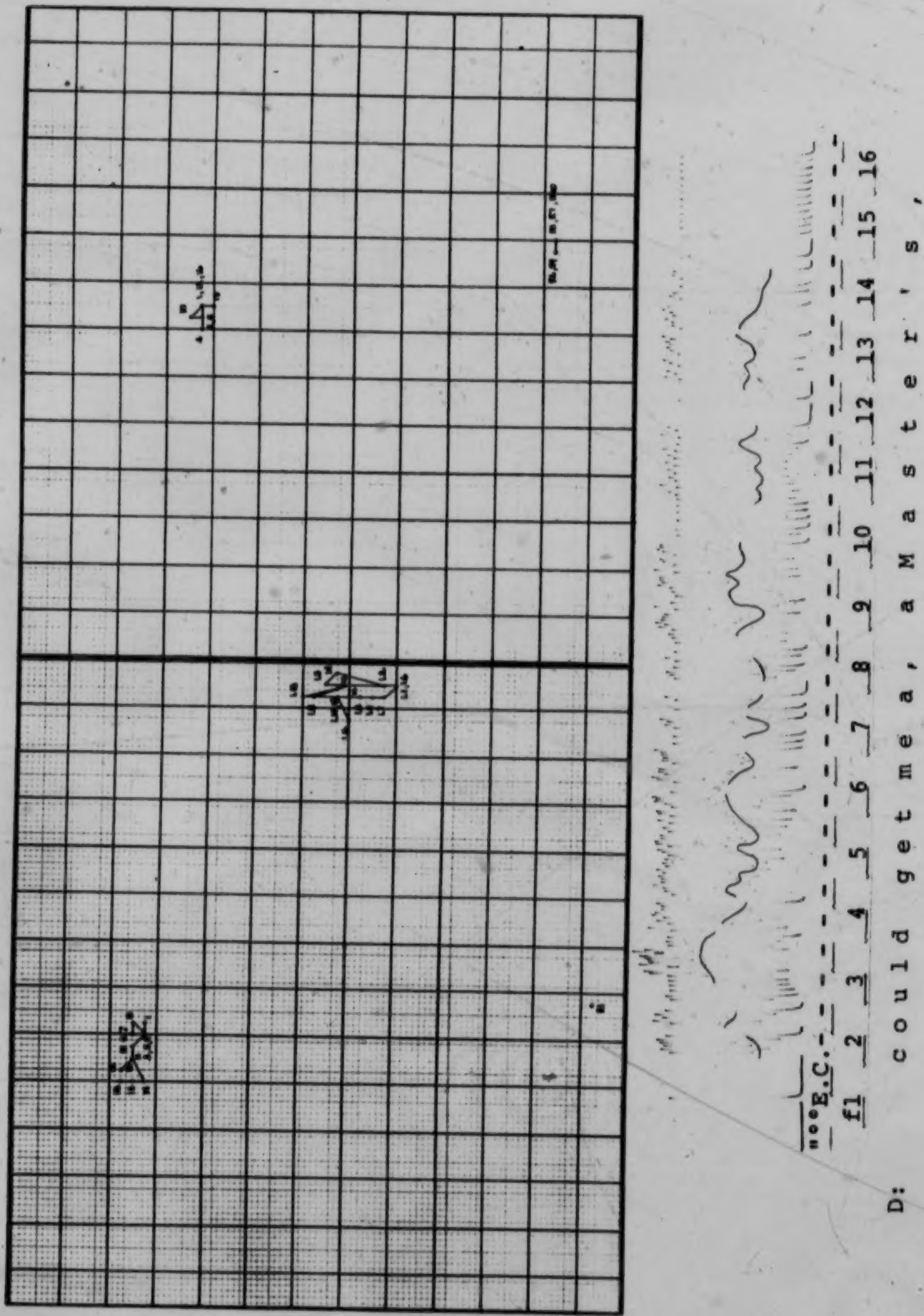
D:	E:
...could get me a Master's	
or even a Ph.D. depending on	
what I built around it.	
In Ed. Psych.	And depending on where
	you did it
In Ed. Psych. [snap]	...In...Oh yes.
Any	
time	Yuh
	Yeah
Here, as a, as a small	Or Ed Admin even
part of the thesis they	
don't want it in.	

(Note: utterances on the same line are produced at the same time, coincidentally. Each new line indicates the next utterance in the sequence.)

Example VI.2

Could get me a, a Master's--is accompanied by an intricate network of head, hand, and face movement. As the utterance begins, D and E are in eye contact. D is sitting

Plate 6.5--Mingograph and Movement Traces: Example VI.2



back in his chair, right arm resting on his chair (off camera), left hand raised, left forearm at roughly 45 degrees, left elbow on chair. E is sitting forward, resting on his elbows, hands and forearms on the table. The way he has them placed, the right arm largely obscures the left, and this continues to be the case for most of this fragment. Movement goes with utterance throughout this small segment. D's movements are particularly obvious. The movement of his left hand is pronounced, and generally follows what might be described as an up-down "wagging" pattern (similar to that characteristic of the stereotypical "Now you listen here!" gesture of the hand). E is relatively motionless. What movement he does make, however, goes in time with the fundamental frequency of D's utterance.

Could--The fundamental frequency rises through most of the word, reaching its apex and turning to fall as the word ends (f2 to 5). D's left hand moves up, angling to the left (L1 to L4). Head movement is confined more strictly to horizontal-vertical axes, going to the right and up (1 to 4). Although not parallel, head and hand movement is simultaneous and follows the direction of motion of the fundamental frequency.

E moves his head right and left horizontally through "could get me a." Here, the maximum displacement of E's head occurs at the same point as the fundamental frequency peaks (4). E makes no measurable hand movement.

Get me a--Fundamental frequency arches through this, rising through "get," peaking in "me," and tumbling over and down with "a" (f5 to 7-8). D's left hand traces a large "V" as it moves down (L5 to L6) and up to the right (L6 to L8). Head movement describes a half square, moving up and to the left (5 to 8). Here, as with "could," head and hand movement occur coincidentally but proceed in opposite directions. And, the change in direction of motion of D's head, hand, and fundamental frequency is coinstantaneous.

E, meanwhile, executes a right-left-right shift with his right hand that goes with D's utterance to the level of the syllable (R5 to R7). Each shift corresponds to and coincides with the production of a new syllable.

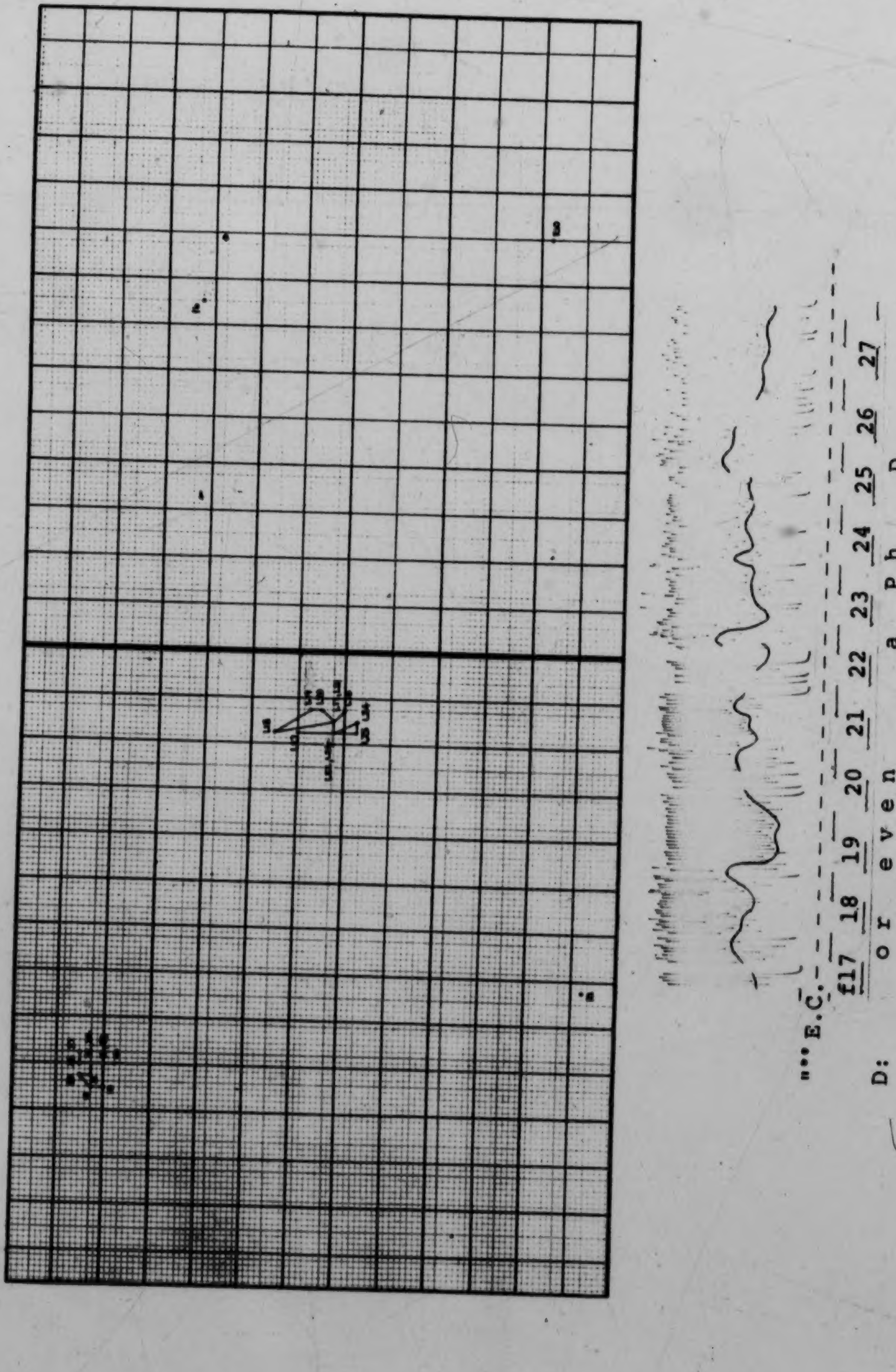
A master's--The intonation curve takes up from where it would have got to had there been no silence. "A" (f8 to 10) is produced with a rising-falling fundamental frequency that follows an arching curve through time. "Master's" (f11 to 14-15) evidences a similar pattern: the fundamental frequency rises through the first syllable and falls on the second. D's left hand rises up with the fundamental frequency (L7 to L11). This movement is part of a series of "jerky" motions--up (L7 to L8); down (L8 to L9); up (L9 to L11)--that seem to imitate the jerky, staccato articulation of "me a, a Mas...". D's head movement here is angular, tracing a capital "V" (8 to 13) across the screen. Change of direction is coincidental, within 0.1-second of the change of direction of the fundamental frequency. The complex of movements that go with the production of "Master's" are of particular note. The transcription indicates that as he produces the first syllable, D moves his head abruptly, to the right and up (10 to 13). Head movement parallels the direction of motion of the fundamental frequency. At the same time, in movements apparent on the video tape but not transcribable here, D opens his eyes wide and turns his head to the right. Succeeding motion of hand and head is exactly parallel, moving downward (12 to 13), and coincident, going with the direction of the fundamental frequency.

E's head movement also goes with, paralleling, the intonation contour. He moves up and down-left through "master's" (9 to 12). His right hand moves horizontally, right-left-right, through "a" (R8 to R10). Right hand and fundamental frequency change direction coinstantaneously here, within 0.1-second of one another.

In the silence between utterances, D's head moves straight down (14 to 16). Coincidentally, D's left hand executes an up-down motion that, going from right to left across the screen, resembles a backwards "N" (L13 to L16). E moves his right hand slightly as well, shifting it up and then down (R14 to R16).

Or even a Ph.D.--D and E continue on in much the same manner through here. There are no major body shifts. D's movement is the more pronounced and obvious. E is in fact motionless for most of the segment. Movement accompanies utterance. A notable series of events not plotted or

Plate 6.6--Mingograph and Movement Traces: Example VI.2-1



plottable on the movement charts begins as D utters the last syllable of this segment. E here begins tiny up and down movements of his head and eyes. This nodding continues for more than a full second, to the point of silence following "built around it" (f44). They are extremely small, precise movements, too small to register on the scale of movement employed in the present study. They apparently go with the fundamental frequency of utterance, and indicate E's agreement with what D is saying.

Or even--The fundamental frequency carries on from "master's": it continues to fall through "or" and rises through "even", completing the S shape of the intonation curve (f17 to 22). D's head moves with the fundamental frequency, paralleling its changings. Specifically, head movement proceeds across-left (17 to 18) with "or," and up to the left, then down or "even" (18 to 21). His left hand describes a similar motion through this portion of the text. That is, the left hand moves up, and down-right (L17 to L21), changing directions with the fundamental frequency.

E is motionless.

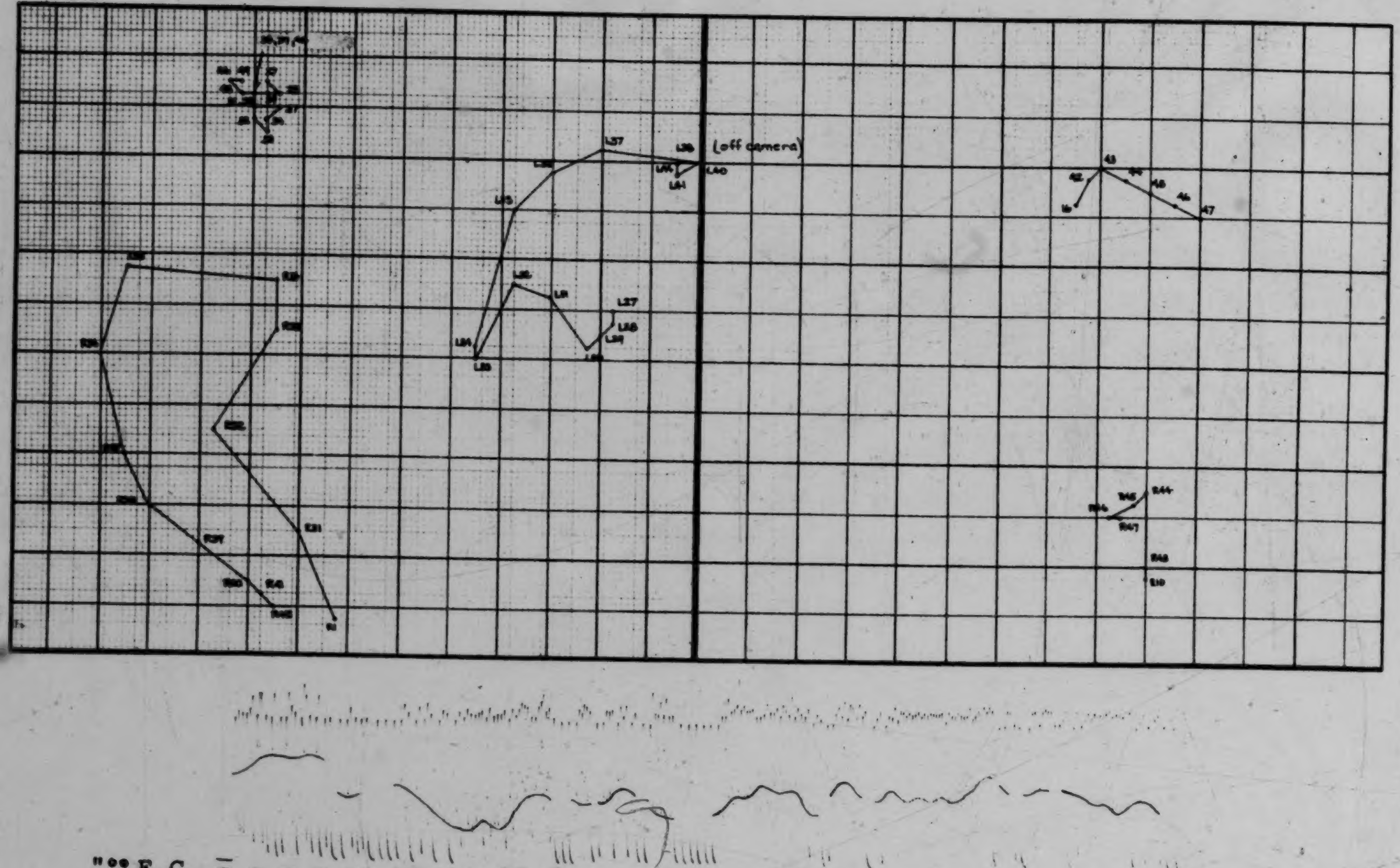
A Ph.D--It is difficult to determine fundamental frequency of this utterance from the mingograph trace. The fundamental frequency would seem to rise to its maximum amplitude and begin to fall during "a" (f22). "Ph.D." (f23 to 28) carries the fundamental frequency through its downward swing and up over the top of the curve. Head and hand movements imitate the pattern of the fundamental frequency, rising and falling with it. D moves his head across-left and down during "a" (21 to 23); and then up, across, and down-left with "Ph.D." (24 to 28). The left hand slips across-right and down to the left on "a" (L22 to L23), and moves across-right and up with "Ph.," slipping down through "D." (L23 to L28).

Depending on what I built around it--is marked by unique hand gestures. On the video tape, D appears actually to build something in the air in front of him as he produces "depending on what I built." Movement goes with speech. This segment also marks the first break in eye contact for the fragment. The break occurs as D constructs the "something" in front of him. His eyes shift away from E to the "something" under construction. E, however, continues to look at D. Eye contact is re-established between them, then, at precisely the moment that D looks back up from his "completed construction."

Once again, D owns the more pronounced movement in the interaction. E is relatively still, except for the very small, unchartable nodding movements described above, and the complex of movement which comes at the close of the segment. These, however, seem to anticipate and be part of the succeeding utterances. Description of them is thus thought to belong more properly to the following segment.

Depending on--It is difficult to ascertain the fundamental frequency of "depending." The mingograph trace is filled with extraneous noise and is particularly

Plate 6.7--Mingograph and Movement Traces: Example VI.2-2



"•• E.C. = - - - - - - - - - ••."

"••E.C. - -

f28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43

D: depending on what I built around it.

illegible. The fundamental frequency apparently climbs and curls over, to begin to drift down through "on" (f32 to 33). D's head movement here consists of a shift across to the right, down, and down-left (28 to 33). He moves down with the fundamental frequency. D's hands begin to come together during this portion of the text (in preparation for the "construction" that is to follow). His left hand describes an arc down and to the right (L23 to L33); his right hand starts up and to the left (R29 to R33). They "meet" and "twirl" in the next 0.1 second, with "what," before proceeding their separate ways.

E produces a chartable nod at the apex of the intonation curve, moving precisely as the fundamental frequency begins to tumble down (31 to 33).

What I built around it--The fundamental frequency arches through "what," rising and falling (f34 to 35). The regular wave-like pattern of the fundamental frequency continues with "I built," rising and falling in steady rhythm (f36 to 38). During "around it" the fundamental frequency moves up and down through time almost onomatopoeically. The contour comes a "round" to round off the top of the fundamental frequency curve (f40 to 41) and moves down through "it" (f42 to 43), proceeding down to the silence at the end of the utterance stream (f43 to 44). D breaks eye contact with E as he completes the "building" in front of him; as the fundamental frequency falls through "built." With "what I built," D's head moves up and down in a saw-tooth pattern, shifting from left to right across the screen (34 to 37). Head movement here parallels the rise and fall of the fundamental frequency. At the same time, with "built," D's right and left hands begin to depart the construction site, heading back to the areas from where they came. Specifically, his left hand moves up and left, slippping off camera (L35 to L38). Half a second later it re-enters the screen area and, coincidentally, with 0.1 second of utterance, comes to rest in a forearm-vertical position (L42), elbow on the arm of his chair. D's right hand slips down-right to the table top (R35 to R43), and in the end--also coincidentally with utterance--achieves a position of rest, palm down, fingers extended, on the table.

The latter portions of this segment are marked by small side-to-side movements of D's head (38 to 43). Head movement reverses direction with the change in direction of the fundamental frequency. E is motionless once again.

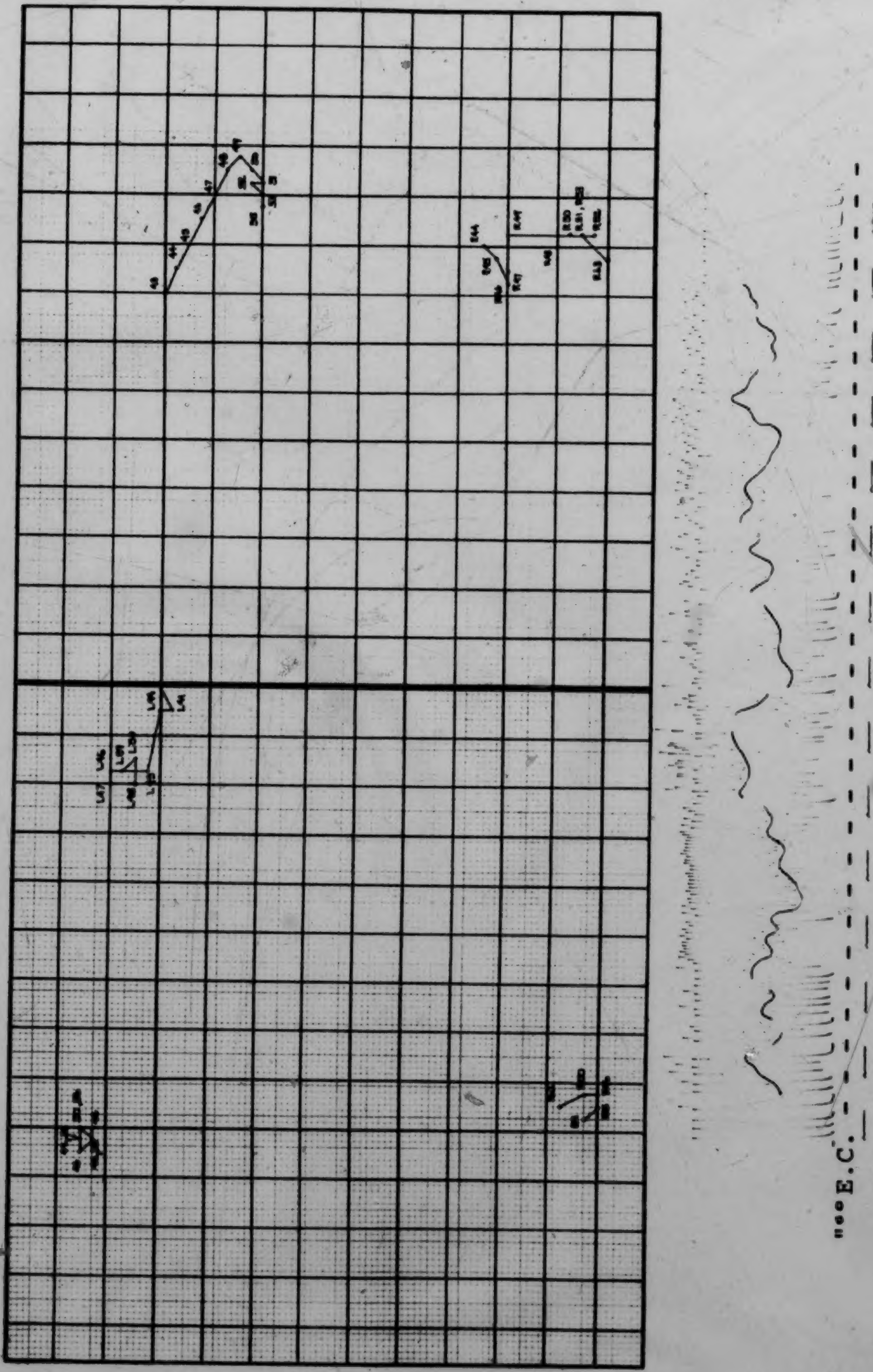
In Ed Psych
And depending on where you did it

Example VI.2a

This segment is a maze of interwoven sound and movement. It is hard to know just where to begin the description. There are no clear-cut beginning and end points. The sound and movement do not cease; everything happens at once. D talks, E talks; D moves, E moves.

E begins this segment 0.2-0.3 seconds before the pause taken as the end point of the previous segment. E sets about raising his head up and to the left, moving toward D, while

Plate 6.8--Mingograph and Movement Traces: Example VI.2a



D: In Ed. Psych.
E: And depending on where you did it.

D completes "around it." As E raises his head, his eyes open wide and he lifts his right hand, extending the index finger until he is literally pointing at D (R42 to R44). With the completion of D's utterance E reaches the apex of the arc described by his head movement and begins a long slide down to the left, toward, or at, D (44 to 49).

And depending--comes in within 0.1 second of the start of his slide and precedes D's "in" by roughly the same amount of time (f44 to 49). Fundamental frequency goes with, paralleling E's head movement, falling in two "bumps" through "depending" (f46 to 49). At the same time, as he slides down, E brings his right hand down to the right, toward D. Index finger extended, he continues to point at D (R44 to R49). Here again, the two-dimensional plot is patently lacking; a third dimension is needed to represent adequately hand movement "as it happened."

In--D's utterance comes in within 0.1-second of E's initiation of "and depending." The fundamentals of the two are so intertwined as to make distinction on the mingograph trace almost impossible. However, from the audio tape, it sounds as though the sustained, almost flat portion of the trace (f45 to 47) corresponds to "in."

D moves very little during this. As noted below, his movements seem to "go with" E's utterance. Nevertheless, they also occur with "in." On the video screen it looks as though D "tenses" his entire body for a split second, "spitting out" the utterance as he releases the tension.

Less poetically, D moves his head to the right and up, away from E, with the latter's "and" (44 to 46), and comes back toward E in time with the falling contour of "depending" (48 to 49). D's right hand hardly moves, remaining at rest, while the left is brought back toward his body, moving slightly downward in two stages with the falling intonation contour (points L45 to L50).

Ed--seems to correspond to the generally rising fundamental frequency on the mingograph trace (f49 to 50). The upward movement documented here agrees with what the ear hears on the tape. D's right hand moves slightly up through the utterance (R49 to R50). Head movement, at this point, and for the most of the remainder of the segment, consists of short side-to-side movements interspersed with short pauses. Pauses seem to last 0.1 second or less; movements have a duration that is usually twice as long.

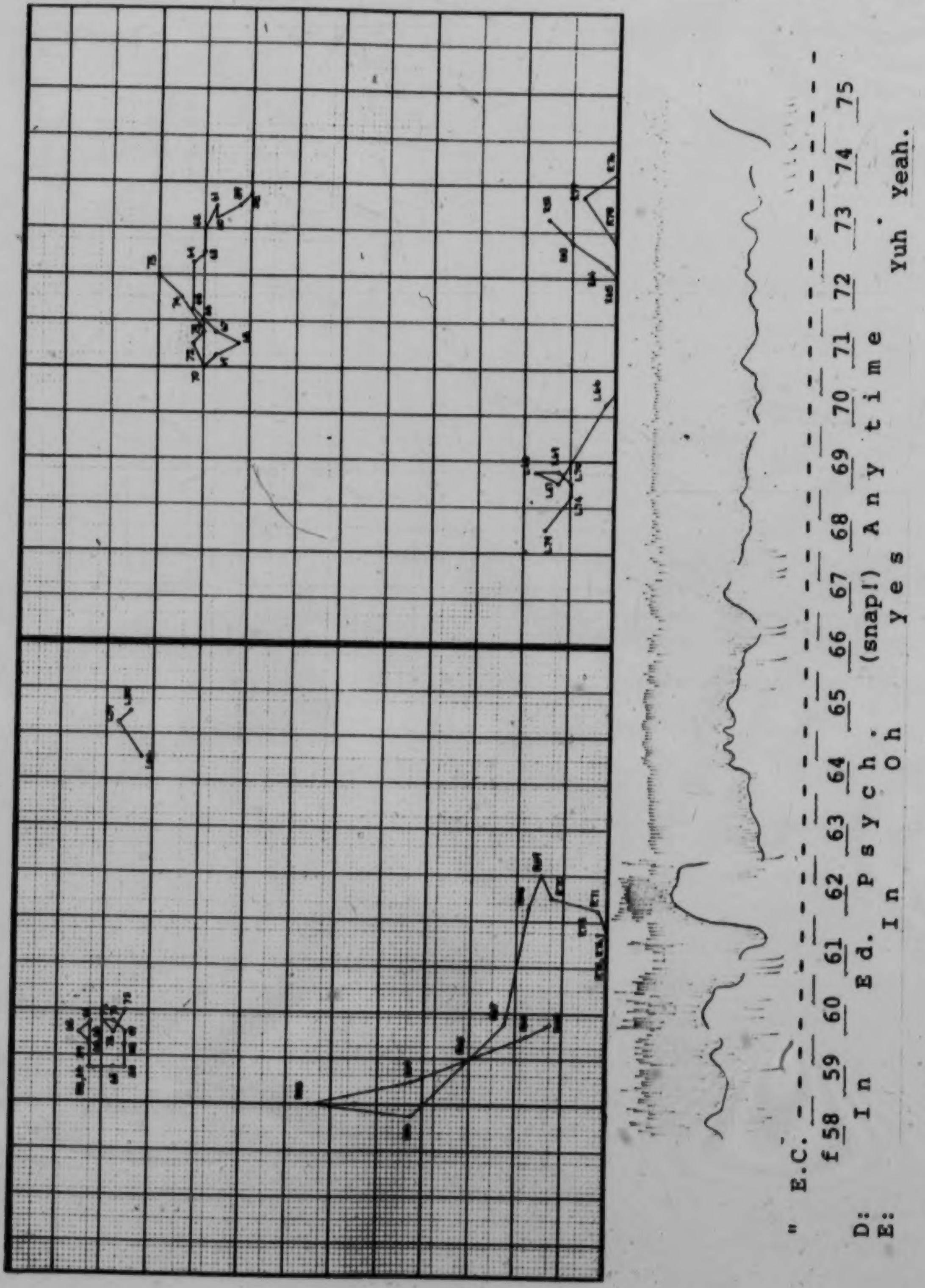
Psych--is a long drawn-out word/syllable, uttered over the last 0.7 second of the segment. It is impossible to attribute the fundamental frequency traces produced by the mingograph here to a specific interactant or utterance. The movement of the fundamental frequency documented in the mingograph traces may correspond to the articulation of E's "on where you did it" (f50 to 57), but the trace is very probably affected by the pitch variation the ear detects on the audio tape through "Psych." In addition, the mingograph record may in fact be the product of a summing process performed on the two acoustic signals by the trans pitchmeter or the mingograph. However, it is possible to say that the fundamentals of both utterances move down as they approach the "silence" that marks the end of the segment (f58). D's head movement goes with this movement of the fundamental frequency : his head slips down and to the right. Because the intonation contour described by the mingograph trace seems to go better with what the ear hears as "on where you did it," the fundamental frequency will be described below.

On where you did it--Overall the fundamental frequency describes a long arch, rising in two smaller arcs through "on" and "where," that peak at the completion of "where (f50-51 to 54-55), and falls off through "you did it" (f54-55 to 57). E's head movement is angular, proceeding down and away from D during "on where" (50 to 54). E's eyes resume a more "normal" character here. At this time he also brings his right hand down and back in toward his body, curling the extended index finger in to his closed hand, and bringing his right forearm to rest against the left (R50 to R53). E is relatively still for the balance of his utterance. D is also quite still through "you did it."

In Ed Psych [snap] Anytime
In oh yes Yuh.Yeah Oh definitely
Example VI.2b

This segment, too, is such a complex interweaving of the sound and movement of both participants that it is very difficult to know where or how to begin the description. Words fail to convey adequately the complexity of the events

Plate 6.9--Mingograph and Movement Traces: Example VI.2b



as they happened. Movement and utterance go together; they are one such separation, even, or perhaps especially, for the purposes of analysis does violence to the reality. E's non-word utterances, portrayed here as "in," "yuh," and "yeah" are particularly difficult to convey in a written text. So much depends on intonation, timbre, amplitude of voice, and the movement that goes with them.

In Ed Psych...[snap]--The fundamental frequency is taken up at the point to where it would have got had there been no silence. It is an emphatic utterance; D is repeating what he has just said.

In--The fundamental frequency first rises slightly, catching the trajectory implied across the silence, and then falls (f58 to 60). D moves his head horizontally, left to right, through the duration of the utterance (58 to 61). Both left and right hands are motionless, still in the "state of repose" achieved with the completion of the previous utterance.

Ed--The fundamental frequency appears to arch through "ed" (f60 to 62). Toward the end of the word the fundamental frequency is obscured, altered by the interjection of E's "in" which also seems to have a rising fundamental frequency but at a much higher frequency. D brings his head up to the right with the fundamental frequency of his utterance, and, at the same time, begins a parallel, though extended, movement of his right hand (62). This hand motion marks the beginning of the sequence that is to culminate 0.5 second later in the "snap" which comes upon completion of "psych."

Psych--D takes up the fundamental frequency at the point to which it would have got had there been no "interruption" by E. The fundamental frequency describes a long arch through here (f63 to 66), peaking just before the utterance ends and slipping down toward the moment where D snaps his fingers (66). The "snap" is noise, and not recordable with the present system. D swings his head down-right and across through "psych" (63 to 66). Head movement goes with, changing direction coinstantaneously with the fundamental frequency. D's right hand continues its upward sweep to the "snap" (R63 to R65).

[Snap!]--D's head and right hand both begin to move downward with the snap (66-67). D and E break eye contact simultaneously at this point. D glances down and away from E, in the general direction of the snap, while E directs his gaze down to the table top in front of him. In the same instant D moves his left hand horizontally to the left (L66).

In...oh yes--E produces this at the same time as D utters "In Ed Psych [snap]." It must be remembered that the description below applies to sound and motion that take place coincidentally and coetaneously with the sound and movement outlined above.

In--This seems to be a partial word, an interjection, the sound of which is most closely approximated by "in." The fundamental frequency arches through very high frequencies with "in" (f62). As previously noted, this interjection trespasses, coming in on top of D's "Ed." E's head moves quickly up and to the right immediately prior to and with production of "in" (59 to 63). He draws away from D, moving back from the forward inclination that went with his (E's) previous utterance. E's left hand is still on the table, off camera and/or obscured by the right arm. His right hand begins moving with the onset of "in," travelling down to the right and off camera (R61 to R65)--eventually coming to rest on the table out of camera range.

Oh yes--are two rather long and "melodious" vocalisations. That is, they come with a "multiple pitch" delivery. Although it is not possible to separate the fundamental of these from that of D's utterance, listening to the audio tape, "oh" seems to follow the rising-falling pattern noted for "Psych" (f64 to 65), and "yes" seems to be found in the succeeding traces of an arching fundamental frequency contour (f66 to 68). E's head continues to move back and away from D here. Specifically, head movement proceeds across and down to the right (63 to 68). On the video screen, E's whole body appears to move forward-right, and down, going with the falling fundamental frequency. E's left hand comes on camera as E produces "yes," moving up and to the right--in toward the body (L65 to L67).

Anytime--has a "rippling" fundamental frequency (f68 to 73). There is an overall rising and falling S shape that is embroidered with the ripples of articulation. It is a pretty bit. Head movement goes with the overall shape of the fundamental frequency. D's head moves across to the left and up with "any..." (68 to 72), coming down and to the right with the fundamental frequency of "time." D's head movement continues past the end of the utterance, and proceeds on down, shifting to the left, toward E (74 to 75). It is here, as he completes "anytime," that D brings his eyes up and re-establishes eye contact. D's right hand continues down from the apex reached with the snap of his fingers, slipping down-left. This motion carries on just beyond the end of the utterance, coming to rest on the table top 0.2 second after the completion of the utterance (68 to 75). D's left hand remains motionless throughout, maintained in the "forearm vertical" position described above.

E also moves with D's "anytime." Head movement follows the contour of the fundamental frequency, moving up to the right with "any..." (68 to 71), and up-left-and-down with "time" (71 to 73). The sharp sawtooth-like movement of E's head across the screen here (71 to 74) also goes well with his own "yuh." On the final rise of the sawtooth pattern (74) E brings his eyes up, in concert with D, establishes eye contact. Finally, E completes the move of his left hand in toward his body here.

Yuh--is a brief, staccato-like interjection that occurs as D rounds off "anytime" (f72 to 74). The mingograph trace reveals a rising fundamental frequency that carries on from the previous (that is, D's) utterance.

Yeah--is uttered alone. D is silent. The fundamental frequency falls through the first syllable and rises with the second (f74 to 75-76). As he initiates the utterance E begins a long, arcing motion with his head. Specifically, he moves up and left. Both E's hands remain motionless through this.

D's head movement, down-left (74 to 76), goes with the fundamental frequency of utterance. Head movement parallels the fundamental through "ye.." and changes direction precisely with the fundamental on "..ah." Tiny movements, like "nodding," embroider the larger movement. The smaller movements are evident on the video tape, but are too fine for transcription here. D's hands do not move.

Oh definitely--The fundamental frequency arches through "oh" (f77 to 78). With the first syllable of "definitely" the fundamental frequency continues to slip down a bit, but turns to arch through the last syllables of the utterance (f78 to 83). The larger arch of the last syllables is embroidered with smaller arches that apparently correspond to syllable production.

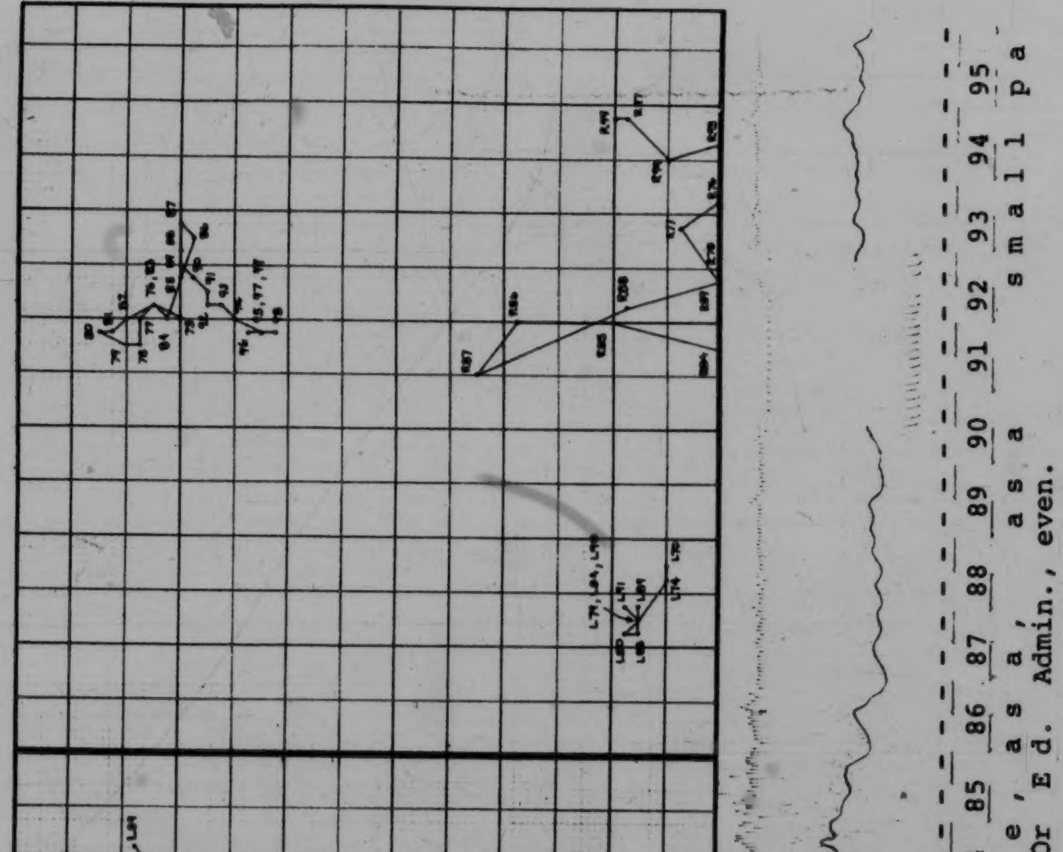
E's head movement goes with the fundamental frequency of his utterance. In particular, head movement parallels the rise and fall of the intonation contour through "definitely" (78 to 83). Sound and movement peak--they reach the apex of their respective arcs and reverse direction--at the same point in time. There is a finer movement involving side to side motion of the head that is obvious on the video tape but is too small to be captured here. Impressionistically, the movement appears to go with syllable production. E lifts his right hand with the rising fundamental frequency of "oh." Right hand reaches maximum displacement with the fundamental frequency, at the centre of the utterance, and drops off camera with the falling fundamental frequency of that word (R76 to R78). E's left hand moves up and right, part of a small body shift that is produced coinstantaneously with the initiation of "definitely" (L78 to L80), and then drops through the last syllables of the word, again, with the fundamental frequency (L82 to L83).

D is motionless.

Here, as a, as a small part of the thesis
they don't want it in.
Or Ed A d min even
Example VI.2c

This segment finishes the fragment, and rounds off topic. Only at the first is there the same degree of sound and movement interweave here that characterises the previous two segments. D speaks last. As he completes his utterance, he pauses, fixing E with his gaze. E looks back, and they share what might be described as a "look of disbelief." And there is a pause: each realises that the other has realised that the other realised that. . .to an infinite regress. It is one of Mair's "timeless moments." Lasting perhaps 0.2 second, it stops the text. Then E shakes his head in disbelief, D recommences his shift to a more upright position in his chair, and the text continues.

Plate 6.10--Mingograph and Movement Traces: Example VI.2c



D breaks eye contact as he produces "here" (83). The transcription indicates only that he moves his head to the right, away from E. From the video tape, however, it is apparent that D turns his head to the right as he moves, shifting his gaze away from E to a spot somewhere in front of him (D). E continues to look at D, but D does not re-establish eye contact until very near the end of the fragment. It is not possible to determine precisely when eye contact is re-established; D's glasses shade his eyes, hiding the direction of his regard. From head and body position and movement it is apparent, though, that D has re-established eye contact as the fragment ends and he and E enter into the "timeless moment" described above.

Here, as a, as a--While "here" is heard on the tape, the mingograph trace documents a fundamental frequency that rises and falls, arching through the word (f83 to 85). The trace where "as a, as a" is heard descends with a "bump" (a short, rising-falling half-circle) through "as" (f86-87) and turns to climb through a double hump curve with "a, as a" (f87 to 90-91) to the brief silence preceding "small." As before, it is difficult to attribute the mingograph trace to a particular utterance through this part of the text. Attribution, then, is somewhat tentative, representing a best possible choice from available data. From the audio tape it seems that the high fundamental frequencies recorded here belong to D's "here." It is an emphatic utterance, anticipating E's just a bit. Mair's model proposes that the high frequencies here, the "pitch excursion," in fact represents D's attempt to regain or to maintain topic control. He goes "over the top" to get said what he wants said. This description goes well with what happens on the tapes. D does ride over E to complete his argument and round off the topic.

D's head movement goes with the overall shape of the intonation contour through "here, as a, as a." Head movement parallels the rise and fall of the fundamental. Specifically, his head moves up to the right and down with "here" (83 to 84). There is a brief pause, and the downward movement continues with "as a" (85 to 88). "As a" is accompanied by a short movement upward (89).

D gesticulates a good deal throughout this segment. It is during this particular section however, that he initiates the movement sequence which culminates slightly more than .5 second later in the complex gesture that goes with, and perhaps virtually constructs, a "small part of the thesis" (at chest height, in the area directly in front of him). In particular, D's left hand veers sharply down and to the right with "here" (L83 to L85) and climbs almost as sharply back up to the left through "as a," (L85 to L87). The second "as a" coincides with a quick "V" movement of the left hand (L88 to L90). The apex of the final stroke of the "V" falls over through a long arching motion which continues down through to the completion of "small part" (L91 to L95). D's left hand changes direction coincidentally with each change of direction of the fundamental frequency.

D's right hand goes with the fundamental frequency as exactly as his left. His right hand sweeps up quickly, angling to the right, with "here" (R83 to R85). There is a smaller downward-right motion with "as a" (R85 to R87) that is immediately reversed in a sharp upward-left gesture (R87 to R88). The second "as a" occurs with this gesture. The gesture continues on into the long angular swooping motion down toward the table with which the right hand closes off the segment.

Or Ed Admin, even--occurs in concert with D's utterance. D and E speak together, in time (rhythm) and in harmony (melody). They begin their respective utterances within 0.2 second of one another. D's precede's E's slightly; D seems to anticipate the beat a bit. (On the tape, judging from the emphasis and tone of his voice, D is determined to get his utterance in before E can enter with his.) "Or Ed Admin, even" follows the intonation contour described above for "Here, as a, as a." The fundamental frequency rises through "or" (f83 to 84), descends with a "bump" through "Ed Ad.." (f84 to 87), and climbs again with "..min, even" (f87 to 90-91).

E's head movement proceeds down-left through "or Ed" (83 to 85). Having moved back and up, away from D, during "oh definitely," E now moves forward and down, toward D. The small side-to-side movements of E's head going with "oh definitely" continue through this shift forward and down.

E's left hand is tucked in, toward his body; it lies on the table in a closed fist position for the balance of the fragment.

E's right hand gestures with the utterance. The shape of the movement parallels the general shape of the intonation contour, rising and falling with it (R83 to R89). As he completes his utterance, his right hand slips down, off camera.

D and E move and speak simultaneously through this portion of the text.

Small part--has a sustained, relatively flat fundamental frequency contour (f92 to 96). The fundamental frequency rises slightly with "sma..." and falls a bit through "...art," but for the most part, is maintained at the level at which the utterance began. D's head movement consists of side-to-side (right to left) motions during "small" (91 to 93), followed by a pause and a downward movement that goes with the falling fundamental frequency of "part" (94 to 96). As already noted, D's hands describe two long arcs through "small part." The arc of the left hand ends with, at the same time as, "part" is completed (L96). Downward motion ceases here. His left hand begins to move up and to the left (L96 to L97).

D's right hand changes direction 0.1 second after the left (R97). Right hand movement effects an upward motion, veers left, and tumbles down off camera (R97 to R98).

E moves with the fundamental frequency of utterance here, as well. E's head angles down to the right through the utterance (91 to 96). This is a generally smooth, sweeping motion, broken at the end by a small "tick" up which coincides with the completion of the /t/ of "part" (96). As D begins "small," E's right hand comes back on camera (R93), moving up to the left in three stages or plateaux. The first coincides with "small," the second with "of," and the third with "the" (R93, R96, R97, respectively). E's right hand comes to rest at this last point and is motionless for the balance of the segment.

Of the thesis--the intonation contour describes a complete sine wave here. "Of"--the fundamental frequency rises slightly (f96 to 97); "the"-- fundamental frequency dips a bit (f97 to 98); "thesis"--the fundamental frequency dips in a more pronounced fashion, and rises as the words end (f99 to 100). D makes a series of short head movements down and to the left through this part of the text (96 to 101). The general flow and direction of the movement is interrupted briefly by a "bump" up, which occurs on the first syllable of "thesis" (99). D's hands angle downward, sweeping off camera toward the arms of his chair. This movement is part of the preparation for a major position shift--he straightens up and sits forward in his chair--that occurs immediately after the completion of this fragment. Specifically, D's left hand climbs sharply through "of" (L95 to L97), swerves right and down with "the" (L97 to L98), and veers left off camera through "thesis" (L98 on). His right hand swoops down to the right, slipping off camera much more quickly than the left hand (R97 on).

E is quite still through this part of the text. There is only one transcribable movement: E dips his head slightly with "of the" (96 to 98). Head movement parallels the rise and fall of the fundamental frequency. But from "of the" to the end of the fragment (and beyond) there is no further discernable movement by E.

They don't want it in--With this utterance D rounds off topic, and coincidentally, completes the fragment.

They don't--picks up the trajectory of the fundamental frequency as it carries on across the instant of silence after "thesis." The fundamental frequency rises slightly on "they" (f101 to 102), then tumbles over and down through "don't" (f102 to 105). D's head movement traces what looks like an awkward "W" on its side, shifting down-left, across-right, down, and up to the right (101 to 104).

Want it--the fundamental frequency continues to fall a bit through the first part of "want," but bottoms out within 0.2 second (f105 to 107-108). Through the balance of "want it" the fundamental frequency varies little, maintaining a constant frequency to "in" (f107-108 to 109-110). There is an animated section of the mingograph trace here that apparently goes with the "noisy" consonants of "don't want" (f104 to 105). The fundamental frequency rises and falls dramatically for a bit, and then comes up to achieve the "maintenance" frequency for the phrase.

D's head movement is equally animated through this portion of the text. Continuing on from the large "W," D's head describes a large arc that follows the rising and falling motion of the fundamental frequency. The larger movement is embroidered with small side-to-side movements that, although apparent in real time on the video tape, are too fine for transcription here. Specifically, transcribed head movement proceeds up to the right (103 to 104), swerves up-left (105 to 106), and then comes down to the left (107 to 108).

In--this, the final word in the fragment, is an extended, "multi-tonal" utterance. Pitch variations are apparent even to the ear. The fundamental frequency arches through, rising with the vowel, sweeping down on the sustained /n/, and coming to rest in silence as D "freezes," caught in the timeless moment. D's head movement parallels the movement of the fundamental frequency, rising and falling in time with it (109 to 111). Just after the fragment ends, D "flips" his head up and toward E. It is a short, quick movement that carries on the implied trajectory of the intonation contour. It is here that D locks gaze with E and becomes "enrapt," moving into the timeless moment, the shared realisation described above.

6.3.2 Sound and Movement Clusters--Interpretation

This rather lengthy fragment is an important one for the present investigation of synchrony in natural conversation. The fragment is filled with examples of coincident, coetaneous manifestations of sound and movement, instances of what are here interpreted as synchrony. It is all there, on the tapes. The middle sections of the fragment, where D and E talk together, at the same time as and on top of one another, are particularly appealing. During these one or two seconds everything happens literally at once and in time with the fundamental frequency of utterance.

There is a mass of data here. In the manner begun in the first fragment, I shall deal with segments separately. Interpretation will take off from the point of view of the model. I shall argue that the selected instances of simultaneously occurring speech and movement within and between interactants in fact constitute evidence for the existence of both intra and interpersonal synchrony as observable, identifiable, and definable events in the process of human communication.

> Could get me a, a Master's or even a Ph.D.,
> depending on what I built around it
>
> Example VI.2d

Could get me a"--The simultaneous movement of D's head and hand, and the movement of both in time with the fundamental frequency of utterance is here taken as evidence of intra personal synchrony in action. Further, such synchronous production of sound and motion within an interactant is thought to be the product of a single central patterning program in the brain. Of particular note in this utterance are the coinstantaneous changes in direction of head, hands and fundamental frequency, and, the parallel movement of all through time. The very small time frame and fine detail of integration argue strongly for the notion of synchrony, obviating claims for a feedback loop or a modified stimulus-response model as generating source for the sound and movement complexes. There is also the "wagging" motion of D's left hand. An obvious movement sequence, it takes on a "semantic" aspect in the interaction. More properly, it forms an integral part of the state of play being delivered with its occurrence.

That E moves in time with D's utterance is argued to support the notion of interpersonal synchrony. It is the precise nature of the timing of his movements with D's utterance that argues for synchrony. With regard to head movement, for example, E achieves his farthest excursion from "home base" at the same time as the fundamental frequency peaks. A tenth of a second later, E begins to shift his right hand back and forth in time with the fundamental frequency of utterance, going with it to the level of syllable production. This, too, is argued to constitute interpersonal synchrony in interaction.

A Master's--D moves with his utterance. The coincidence of sound and motion within a single interactant is interpreted as intrapersonal synchrony. Specifically, D's left hand parallels the movement of the rising fundamental frequency through "a," and, on a larger scale, imitates with jerky head and hand motions the staccato articulation of "me a, a Master's." Furthermore, in a separate movement sequence, D's head changes direction of motion coinstantaneously with the change of direction of the fundamental frequency. Again, movement follows utterance much too quickly and with too much precision to be part, the product, of a feedback system.

The complex movement of head, hand, and eyes produced with "Master's" is an outstanding example of what is here taken to be intrapersonal synchrony. It is a simultaneous, multimodal production. D's head shifts back and turns to the right; his eyes widen; his left hand gestures; the word is uttered. The time frame and degree of integration of the two modalities argue against the stimulus-response model as generating source. The notion of intrapersonal synchrony, simultaneous manifestation of sound and motion through a single central patterning program in the brain, would seem to account most adequately and accurately for what is perceived. Sound and movement occur together, as one, integral parts of the state of play being delivered with their occurrence. D's subsequent motions of head and hand, paralleling not only one another but the movement of the fundamental frequency, are taken as further evidence of intrapersonal synchrony in this portion of the text.

E moves with D, and with the fundamental frequency of utterance. There is conspicuous coordination between the two that is beyond the level of fortuitous coincidence. The simultaneous manifestation of sound and motion between interactants is intepreted as interpersonal synchrony. E's right hand moves with the fundamental frequency of "a," changing direction precisely in time with the fundamental frequency. His head movement parallels, going with, the movement of the fundamental frequency through "Master's."

The silence between utterances--that is, between "Master's" and "or even"--presents a graphic precision of movement between D and E. They initiate movement together, within 0.1 second of one another. Again, the time frame and precise integration of modalities argue strongly for the concept of synchrony.

Or even the Ph.D.--D's head and left hand describe trajectories which are coincidental and congruent with that of the fundamental frequency contour through "or even a Ph.D." Sound and movement proceed together through and in time. They are thought to be the product of a single central patterning program in the brain and are here taken as evidence of intrapersonal synchrony.

Depending on what I built around it--The hand gestures which accompany this utterance have been described as a "construction": the hands appear to be building something around the questionnaire as D produces the utterance about it. The coincidence of sound and motion is argued to strongly support the concept of intrapersonal synchrony. Sound and movement, the visual and aural modes, are working together to deliver the state of play; hand movement is not merely an addendum to, or an illustrator of, what is being said. The very small time frame involved does not allow for such attribution. Furthermore, D's hands move coincidentally and coetaneously with both head movement and intonation contour.

D's head movements follow the trajectory of the fundamental frequency through "on what I built around it." Sound and movement peak and reverse directions coinstantaneously. Sound and movement also cease simultaneously. Within 0.1 second, D completes his utterance and brings his hands down to rest on the table top. This close imitation of aural and visual modes within one interactant is interpreted as intrapersonal synchrony, and is, in addition, thought to be the multimodal product of a single central patterning program in the brain. Finally, it is interesting to note that D's production of "around it" has a fundamental frequency that in fact does come up and "around it." That the fundamental frequency takes on the onomatopoeic character may constitute further evidence of intrapersonal synchrony, but there is not enough data to tell. For now, then, it remains a remarkable coincidence of sound and meaning.

In Ed Psych
And depending on where you did it

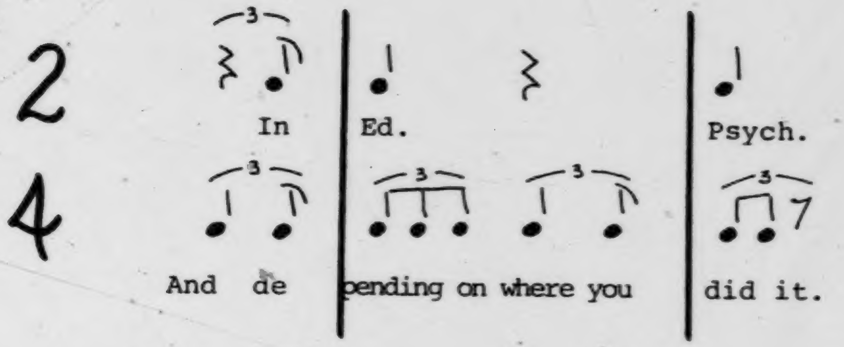
Example VI.2a

D and E "come in" together here. This simultaneous initiation of utterance is interpreted as evidence of interpersonal synchrony. The sharing and interweave of speech rhythms through the duration of the utterance is taken to be a manifestation of interpersonal synchrony on a grander scale. Both interactants move and speak together, and, it is argued, in time to an underlying supra-individual "beat." Figure 6.4 is a demonstration, using standard Western music notation, of the rhythm of this portion of the text in which D and E vocalize together.

The movements of both D and E also go with the fundamental frequency, proceeding in and through time with the melodic pulse of utterance created by the S-shaped falls of the fundamental frequency. The coincidence of sound and motion within and between interactants is interpreted as synchrony as well. Indeed this entire segment is argued to be complex manifestation of the several and various forms of synchrony.

In--D freezes momentarily and "unleashes" the utterance. The "explosion" produces sound and movement, although much of the latter is too fine to be transcribed here. The tiny explosion of sound and movement in one interactant is argued to constitute evidence of

Figure 6.4--Musical Notation of "And depending.."



intrapersonal synchrony, and is further thought to be the product of a single central patterning program in the brain.

And depending--E begins a complex of movement just prior to this utterance. He raises his head and right hand, leans forward, and opens his eyes wide. The complex of movement goes with the fundamental frequency contour of D's utterance. That is, E initiates the upward motion of his head as the fundamental frequency rises through "around it." Precisely as D completes the utterance, E reaches the apex of the arc described by the movement of his head. Such coincident, coetaneous production of sound and motion between interactants is taken as evidence of interpersonal synchrony in action. The precise integration of modalities between persons and the very small time frame demand explication in terms other than those proposed by a stimulus-response model. It is far too quick and too detailed to be thus explained.

Mair's model proposes that a falling fundamental frequency delivers the immediate state of play of which it is a part to a stable state (from which the interaction may take new direction). It might be argued, then, that the falling fundamental frequency of "around it" cues E's sequence of movements. Although a fascinating possibility, such speculation is beyond the scope of the present study. However, it is argued that the smooth sequencing of events and fine interlace of repartee may be taken as evidence of interpersonal synchrony in action. In addition, it is proposed that the synchronous co-patterning of sound and movement among interactants suggests that they should be considered as nodes of a supra-individual system in evolution.

E brings his head down and toward D, raising his right hand to point at him (D), as he (E) produces "and depending." The arc described by the fundamental frequency and the movement of E's head follow each other down. This coincidental realisation of sound and movement within an interactant is thought to support the notion of intrapersonal synchrony.

Throughout this segment, utterance and movement overlap, interweave, and come in on top of one another. It is therefore not always possible to attribute the mingograph record of fundamental frequency to a specific individual with any degree of certainty. The uncertainty affects primarily interpretation of an event as intra- or interpersonal synchrony. As a result, in the discussion that follows, "type" of synchrony is not always specified or specifiable. "Synchrony" will be employed as the general term in these instances.

Ed--Through "Ed" there is a generally rising fundamental frequency, accompanied by an upward motion of D's right hand. The coincidence of audio and visual events is interpreted as evidence of synchrony.

Psych--In the final 0.1 second of this utterance, fundamental frequency slips downward into silence. D lowers his head slightly in time with the falling fundamental frequency. That the two modalities change direction simultaneously is taken as an indication of synchrony.

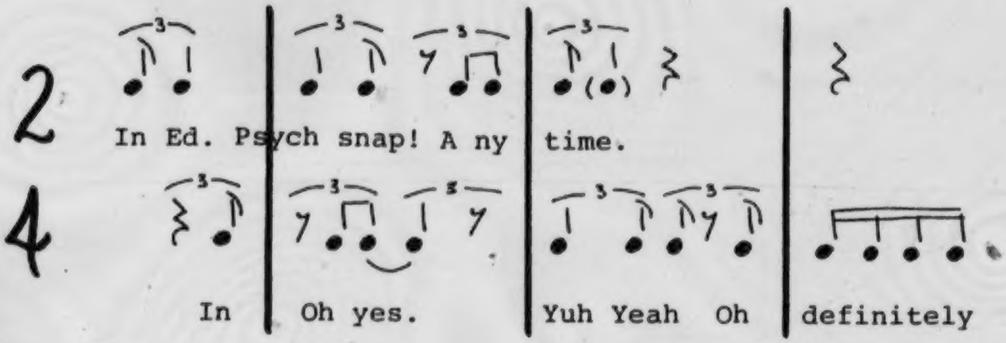
On where you did it--The fundamental frequency describes a long arc through here, peaking in "where," and falling over through the remainder of the phrase. As the fundamental falls both D and E become quite still, assuming positions of repose. From the point of view of the theory, this simultaneous diminution of arousal of both interactants in time with the falling S shape of the intonation contour is interpreted as marking an instance of synchrony; that is, the shared, synchronous realisation of the stable state of affairs that comes with the completion of the downward slope of the intonation contour, and often, as in this case, accompanies harmonious topic conclusion. With the present example, there is, at best, a momentary stability.

```
In Ed Psych  [snap]        Anytime
In   oh    yes               Yuh.Yeah  Oh  definitely
                                      Example VI.2b
```

This segment is much like the previous in that it is composed of a more or less continuous series of instances of simultaneous and coetaneous sound and movement within and between interactants--what are here interpreted as the several and various forms of synchrony. There is more contrapuntal interweave of aural and visual modes here, though. D and E tend to come in right on top of one another less frequently. Movement and voice tend to fade in and out with less overlap. This "smooth sequencing of repartee" between interactants is strongly indicative of interpersonal synchrony in action. It is proposed that interactants are moving through time in time with one another, with the supra-individual rhythm which they have themselves negotiated and/or created. Again, the rhythm is demonstrable with Western music notation, as in Figure 6.5.

<u>In</u>--D executes a left-right horizontal movement of his head here. It occurs at and in the same time as his utterance. This simultaneous manifestation of speech and movement within one interactant is argued to support the notion of intrapersonal synchrony. Further, it is proposed to be the product of a single central patterning system in the brain.

Figure 6.5--Musical Notation of "In Ed Psych [snap]"



Ed--D brings his head up-right and begins an upward sweep with his right hand as he produces "Ed." These coinstantaneous movements go with, paralleling the direction of movement of the fundamental frequency. As before, the coincidental nature of the production of sound and movement in one individual is thought to support the notion of intrapersonal synchrony.

Psych--D's head swing down-right and across goes with the fundamental frequency. Furthermore, sound and movement change direction simultaneously, in the same 0.1 second. That the motions are coincident and parallel is argued to be evidence of intrapersonal synchrony in interaction. At the same time, such close coordination of sound and motion argues strongly for the notion of a single central patterning program in the brain as "source" for the sound and movement. Time constraints and detail of coordination between modalities are such that more traditional explanations based on the feedback model are simply inadequate.

[Snap]--The moment of the snap itself is one in which sound occurs simultaneously with the movements of D's head, hand, and eye. This coinstantaneous production of sound and motion is interpreted as intrapersonal synchrony. D and E break eye contact simultaneously with the sound of the snap. Both look down and away. It all happens in the rush of sound and motion. It is the precise integration of, between, interactants and sensory modalities within a very small time frame--0.1 second--that argues so strongly for the notion of synchrony.

In--E's "in" is an explosive utterance. It has a high fundamental frequency and is accompanied by a coincidence of sudden movements. As E produces "in" his head and whole body shift quickly up and away from D. At the same time he moves his right hand down toward the table top. There is a flash of eyes--E appears to look closely at D. It is all over in an instant. It is an emphatic, jerky, complex of motion that goes well with the range and contour of the fundamental frequency. This coincidence of sound and motion within an interactant is interpreted as intrapersonal synchrony.

E freezes for an instant immediately after "in." The text continues. On the video tape it looks as though E is struck with the realisation of what D has just said, and is about to say again--or perhaps he (E) is about to add more new information to the discussion. Regardless, it is apparent that E does grasp the meaning of D's utterance long before it is complete. From this point--from saying "in"--E begins to nod, motioning his agreement with what D is saying as the utterance proceeds. Such communal cognitive processing suggests synchrony and, further, hints that interactants may be considered as moving together in and through time, nodes of a single system in evolution.

Oh yes--this is a long utterance accompanied by and occurring with a smooth arcing motion of head and body. Physical motion goes to the right and down at the same time as the intonation contour drops. The intensity of E's utterance, judged from the audio recording, is matched by the emphatic nature of the gestures with which he produces it. The coincident, congruent nature of this complex of sound and motion is interpreted as intrapersonal synchrony.

Anytime--The simultaneous and corresponding movements of D's head and fundamental frequency through here are argued to constitute intrapersonal synchrony in action. The agreement of shape and directionality of sound and movement are the keys.

Interpersonal synchrony is argued to be evident in the coincidental and congruent character of the motion of E's head and the fundamental frequency. That is, E makes a series of small head movements that go with the movement of the fundamental frequency of D's utterance. Sound and movement shift through time together.

Yuh, Yeah--E inserts "yuh" into the interaction in time with D's "anytime." He matches fundamental frequency as well. From the audio tape, it is also apparent that the two utterances round off simultaneously, within 0.1 second of one another. The model proposes that D and E are using an awareness of the underlying rhythm of beat of interaction to perform appropriate interactive behaviours--in this case they close off utterance coinstantaneously. Such smooth production of coincident, coetaneous events between interactants is interpreted as interpersonal synchrony. That D and E simultaneously seek and re-establish eye contact during this utterance is taken as further evidence of interpersonal synchrony. D and E move together, in concert, here.

With the initiation of "yeah," E moves his head upward, beginning a long arcing motion. His head movement goes very nicely with, paralleling the rising fundamental frequency with which the utterance ends. D's head movement also goes with the fundamental frequency of E's "yeah." His head first parallels and then changes direction coinstantaneously with (within 0.1 second of) the fundamental frequency. D,E and the fundamental frequency move as one through this. The coincidence of speech and movement within and between interactants is argued to be in fact evidence of intra and interpersonal synchrony in interaction.

Oh definitely--This utterance is accompanied by several sets of movement that are interpreted as intrapersonal synchrony, manifestations of a single central patterning program in the brain. First, the arc described by the motion of E's head is coincident and congruent with that of the fundamental frequency contour. Second, the finer side-to-side motion of E's head, although too fine for transcription here, is thought to be synchronous with the fundamental frequency of utterance down to the level of the syllable. Third, E raises his right hand with the rising fundamental frequency of "oh." Both hand and voice reach the apex of their respective trajectories at the same point in time, dropping off together. Finally hand motion (both hands) is directed down with the falling fundamental frequency through "definitely."

There is nothing that could be interpreted as interpersonal synchrony here. D is silent and motionless.

Here, as a, as a small part of the thesis
they don't want it in.
Or Ed A-d-min even

Example VI.2c

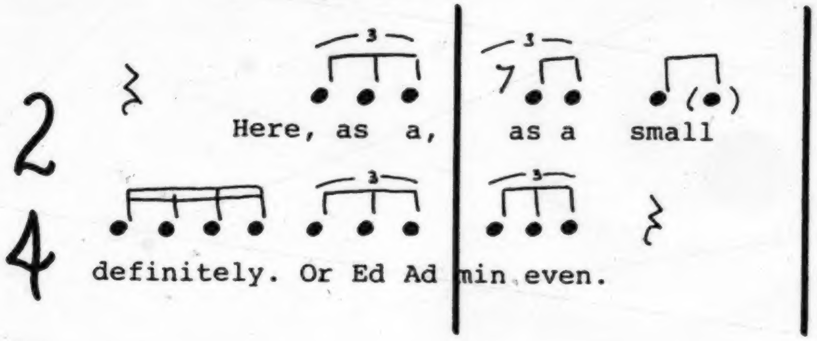
The first moments of this segment continue the intricate interlace of sound and movement. Emanating from both participants, the coincident, coetaneous production of sound and movement strongly support the notion of synchrony. Following the initial burst of activity, the segment is relatively quiet. D speaks alone, and there is not much movement. It is proposed that during the latter portions of this segment, E has given over control to D, allowing him the opportunity to round off topic and complete this section of the text. D takes the opportunity and succeeds in style. The success of his venture is evident in the "timeless moment," the shared realisation to infinite regress, with which this fragment ends.

At the conclusion of his utterance, D completes the argument, summarizing their complaint about the decision of D's dissertation committee. As he finishes the cognitive model, D brings the intonation contour down through a falling S shape, rendering the physical model stable. He brings E with him to the same cognitive place at the same point in time. D has done it; they are there, together in time.

Here, as a, as a and or Ed Admin, even--These begin and end within 0.1 second of one another. This simultaneous initiation of utterance, coming in together, is taken to be a manifestation of interpersonal synchrony in action. D and E move and speak in time with one another, following a supra-individual "beat." Head and hand movement of both goes with the melody, and thus the rhythm, of the fundamental frequency of utterance. The melodic divide of utterance is reproduced in the mingograph traces. Figure 6.6 approximates the rhythmic divide of utterance of this sequence in Western music notation.

Here, as a, as a--D's head movement traces an arc that imitates the general shape of the fundamental frequency contour through this utterance, rising and falling with it. More specifically, head movement and the fundamental frequency change direction coinstantaneously on two separate occasions: first as the fundamental frequency peaks in

Figure 6.6--Musical Notation of "Here, as a..."



"here"; and second, as the fundamental frequency bottoms out in "as a." D's hand movement goes with the fundamental frequency even more exactly. Both left and right hands change direction coincidentally with each and every peak, positive and negative, of the fundamental frequency. It is a detailed congruence of sound and motion. Furthermore, there is a coincident initiation of movement and utterance "here." The complicated series of gestures that accompany/go with a later portion of text, "small part," originate here, precisely as D begins "here." Finally, with all this, it is in this 0.1 second that D breaks eye contact, shifting his gaze to his right and away from E. The entire constellation of coincident, coetaneous sound and movement is interpreted as intrapersonal synchrony.

Or Ed Admin, even--E executives a forward and down movement of head and body as the fundamental frequency travels down through "Ed Ad...". His right hand gestures with the utterance. The direction of motion is the same for hand and fundamental frequency movement. The precise integration of modalities that is evident in these events is argued to support the notion of intrapersonal synchrony.

Small part--As the fundamental frequency falls on "..rt," D effects a downward head movement. Hand movement also proceed downward, falling with the fundamental frequency through "part." Head and hand change direction abruptly, reversing precisely as the utterance ends. What is more, during articulation D's hands make gestures that are thought to go with utterance in that they form part of the state being delivered with their occurrence. D twirls and spins his hands, performing movements which appear to construct a virtual "small part" of the thesis directly in front of him.

E moves with the intonation contour here, as well. His right hand moves upward in time with the rising fundamental frequency of "small." He interrupts a smooth down-right arc of his head with an upward "jitter," just as "part" is completed. The close integration of the movement of one with the utterance of the other is argued to provide evidence in support of the notion of interpersonal synchrony.

Of the thesis--E dips his head slightly, moving in time with and imitating the direction of the fundamental frequency of D's speech. That the movement in aural and visual modes is coincident and congruent between interactants is thought to exemplify interpersonal synchrony.

They don't want it in--There is no evidence of interpersonal synchrony during this portion of the text. E does not move or speak. He remains motionless and silent, looking at D. There are, however, instances of what is interpreted as intrapersonal synchrony. D moves with his utterance. Head movement describes a large arc, following the overall rise and fall the fundamental frequency. More specifically, the fundamental frequency falls abruptly with "don't" and rises quickly through "want." D's head parallels these smaller up-down motions of the fundamental frequency as well, going with them in time. At the same time, D shakes his head from side to side, apparently manifesting with and in motion the negative meaning of his words. The final word of the fragment, "in," provides a last instance of what is interpreted as intrapersonal synchrony. Head and fundamental frequency slip down through time together into silence.

6.4 Example 3

This 8.4-second fragment of text occurs approximately five minutes into the recorded portion of a conversation between two females, C and L. They have not seen each other for some time, and they are bringing each other up to date on their respective activities during the past weeks. As the tape begins, C is outlining her new morning routine. Almost as an aside she explains that, for a number of reasons, dropping her daughter off at the babysitter's in Highlands (a district in Edmonton) tends to intrude on her university work schedule. Talk turns to discussion of favourite areas of the city. C and L compare notes on most preferred areas. C admits that although she considers the area in which she now lives to be "really neat," she prefers Highlands. L agrees with C's choice of neighborhood. L herself lives in Highlands area and remarks that she has always found it to be one of her favourite areas in Edmonton.

There follows a more detailed comparison of the area where C lives with the Highlands area. This comparison in turn leads into a discussion of the type of schools found in these areas. L observes that although the area along Ada Boulevard and 111th Avenue may be a relatively exclusive area, the schools in Highlands are presented with students from various social and economic backgrounds, in large part because of the "feed-in" from other parts of the Highlands area, many of which are lower income. L remarks that the mix does not seem to be at all deleterious for those involved: "[it] didn't seem to bother us when we were goin' to school." C agrees, and then wonders aloud if she has ever been a replacement teacher at Highlands Junior High. She concludes that she is confusing Highlands and another school, a vocational junior high, with a similar name.

As the selected fragment begins, C is relating the antics perpetrated by some of the students at the vocational junior high of interest when C worked there as a substitute teacher. C remembers that she had been told that she wasn't supposed to be rough on the kids, so she was not, at first.

The students showed their appreciation by pulling a clever prank: they turned the clock ahead one half hour and got out of school a half hour early. After that episode, C decided to alter radically her behaviour--the fragment is self-explanatory after this point.

Following this fragment, C and L go on to a very brief discussion of C's experiences as a replacement teacher. L then shifts the topic abruptly. She enquires after C's daughter: "How's B doing?", and the text continues into the new topic.

This fragment was selected for analysis for the numerous instances of intra- and interpersonal synchrony manifest in the audio and video recordings of this portion of the text. For much of the interaction outside this fragment it is apparent that, in terms of amount of talk and movement, C has by far the largest portion of the interaction. The paucity of movement evidenced by L during this half hour of conversation clearly demonstrates the futility of searching for invariant relationships between utterance and movement. In this fragment, however, C and L move and speak together in time. Voices interweave in smooth, closely timed sequence. Movement goes with the fundamental frequency of utterance within and between interactants. In addition, there are examples of finely choreographed series of movements between the two interactants.

The topic is basically L's. She has returned the text to a subject of discussion before taping began. It is C, however, who sets the rhythm for the fragment in the lengthy speech with movement stream which immediately precedes the selected portion of text. And both interactants follow the rhythmical pulse of utterance in time as they develop topic. It is proposed that, through this fragment, C and L are in a state of rapport. They are thought to be literally immersed in the shared text of the interaction.

6.4.1 Sound and Movement Clusters--Description

As the fragment begins, C is sitting up in her chair. Her right hand is resting on the table top. Her left hand is placed on the side of her neck, just under her chin. Left elbow rests on the arm of her chair. C has turned in her chair so that she is directly facing L. Although it is not clear, it appears that C is leaning on her left elbow here--her upper body is inclined to the left, toward L. L is sitting up and forward in her chair. Both forearms rest flat on the table top. Her hands meet in front of her, on the table, right hand on top of left. L has also turned a bit in her chair so that she too is facing her conversational partner.

6.4.2 Text of the Selected Fragment

C:	L:
You know,	[laugh]
And after that I thought	

'You little creeps!'	
[ha-huh!]	[ha-huh!]
So I was	
	[ha!]
[ha!]	[ha!]
	They aren't so slow!
That's right! So I was a real old bitch, and they behaved themselves, and they did some work!	

Example VI.3

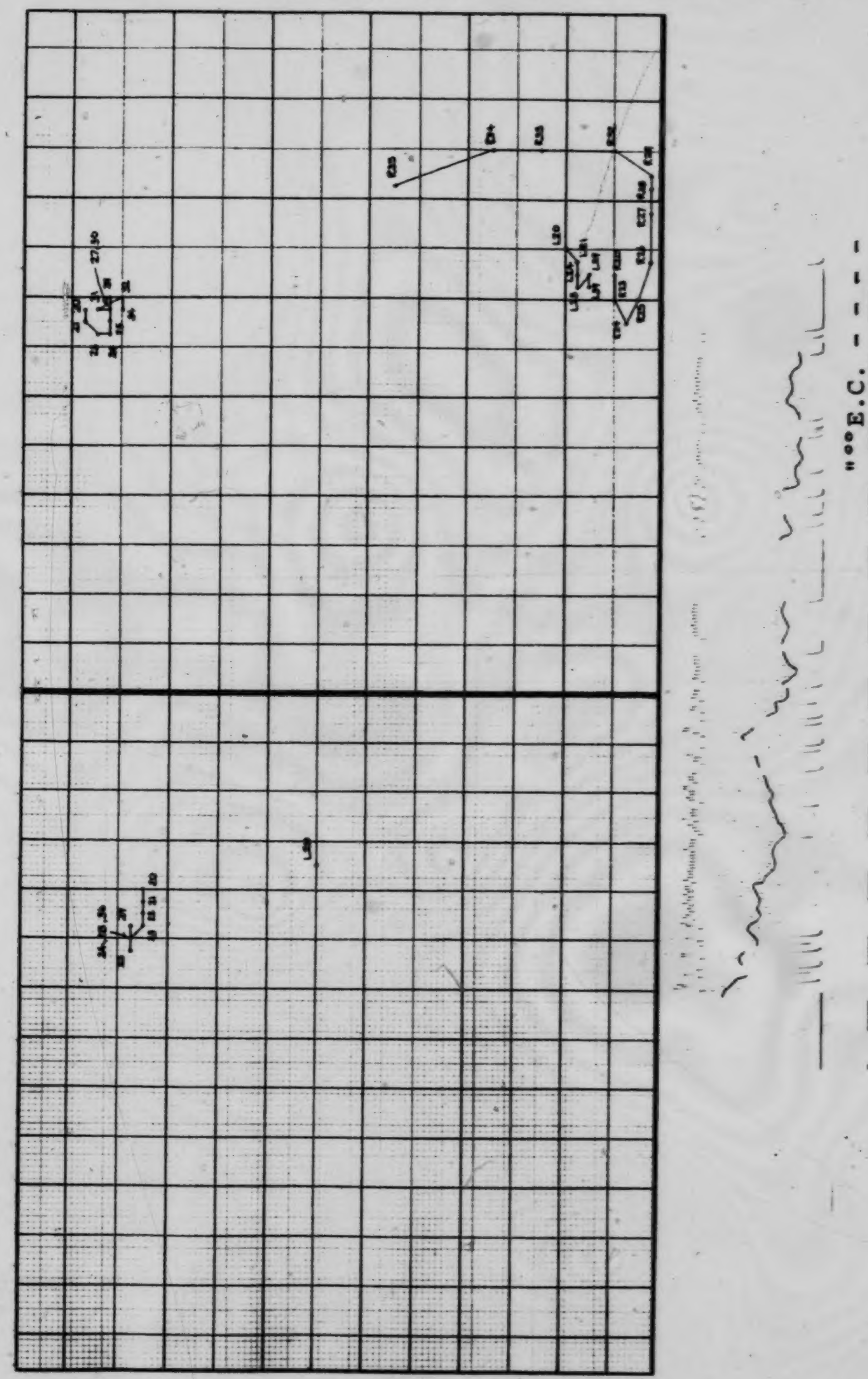
You know, and after that I thought
[laugh------------]

Example VI.3a

[laugh]--It is not possible to determine a fundamental frequency for this utterance from the data presented on the mingograph trace. The utterance simply does not register legibly. From the video tape, however, it is apparent that L does move with the pulses of her laugh. As she produces her first "ha," her head moves up and back (20 to 21), and her left hand moves down (away from her body) along the tabletop (L20 to L21). On the second pulse or "ha," she brings her head down and right, moving it in toward her body and away from C (22 to 23), and shifts her right hand to the right along the table top (R22 to R23). As she lowers her head (22), L breaks eye contact, looking down toward her styrofoam cup.

C breaks eye contact coinstantaneously: she looks down toward the table (22). The balance of C's head movements in this portion of text go with L's utterance. C moves across-right with the first pulse of L's laugh (20 to 22).

Plate 6.11--Mingograph and Movement Traces: Example VI.3a



C: "°°E.C. - - -

L: f21 22 23 24 25 26 27 28 29 30 31 32

[laugh Y o u k n o w], and after that I thought

At the same time, her right hand--off camera--moves in the opposite direction, flipping over from side- to palm-down on the table top. C's left hand remains motionless, resting on her neck just under her chin.

You know--the fundamental frequency rises through "you" (f22 to 23), turns and falls with the first bit of "know" (f23 to 24), and bumps back up as the utterance ends (f24 to 25). This utterance comes in within 0.1 second, almost on top of L's second laugh. On the audio tape they tend to merge and blend into a single unit of sound. C continues the head swing across to the right (away from L) that she began above. Head movement goes with fundamental frequency. As the fundamental frequency falls through "know," C's head shifts up-right (23 to 24). It is a quick shift to a new level or plateau that does not interfere with the progress of the horizontal motion.

L also moves in time with utterance. Her head moves down with the falling fundamental frequency of "know," (23 to 24), and swings across-left, toward C, as the fundamental frequency rises and the utterance ends (24 to 25). At the same time, through "know," L's right hand begins to slip left across the table to her cup (R23 to R25). At each change in direction of the fundamental frequency there is a concomitant, coincident direction shift in right hand movement. L's right hand moves right with "you," swerves down with the falling fundamental frequency of "know," and shifts down-left as the fundamental frequency turns and rises. L's left hand moves steadily to the right through the entire utterance (L23 to L25).

And--the fundamental frequency continues to rise, arches over, and moves down through the end of the word (f25 to 26). The single transcribable movement is made by L. She continues to reach for her cup; her right hand moves left and down (away from her body) a bit (R25 to R26).

After that--the fundamental frequency describes a long arc through here, rising with "after" and turning to fall with "that" (f27 to 29). A good bit of the trajectory of the intonation contour here is implied across the silence recorded by the mingograph between voiced segments of utterance. C moves her head back to the left through and with the long arc described by the fundamental frequency (27 to 29). She turns her head back to the left to where she is facing L once again. At the same time, L swings her head to the left so that she too is once more facing C (26 to 29). Eye contact is simultaneously sought and re-established as the fundamental frequency slips down through "that I" (f29 to 30). L maintains the direction of her regard, looking at C, for almost four seconds--until she completes her own utterance and, glancing down at her cup, prepares at last to have her sip of tea. In contrast, as described in more detail below, C shifts her gaze frequently during these four seconds.

L continues to reach for her cup through "after that," right hand sliding left across the table (R27 to R29). L's left hand makes one brief movement, down-left (L28 to L29). It is coincidental with and parallels the change of direction of the fundamental frequency through "that."

I thought--As the fundamental frequency falls slightly through "I" (f29 to 30), L moves her head to the right a bit, away from C (29 to 30). The intonation contour is relatively flat through "thought" (f30 to 31-32). The slight arch evident in the middle of the word, however, is coinstantaneously mirrored in the up-down movement of L's head (30 to 32). She moves with the utterance.

C is motionless here.

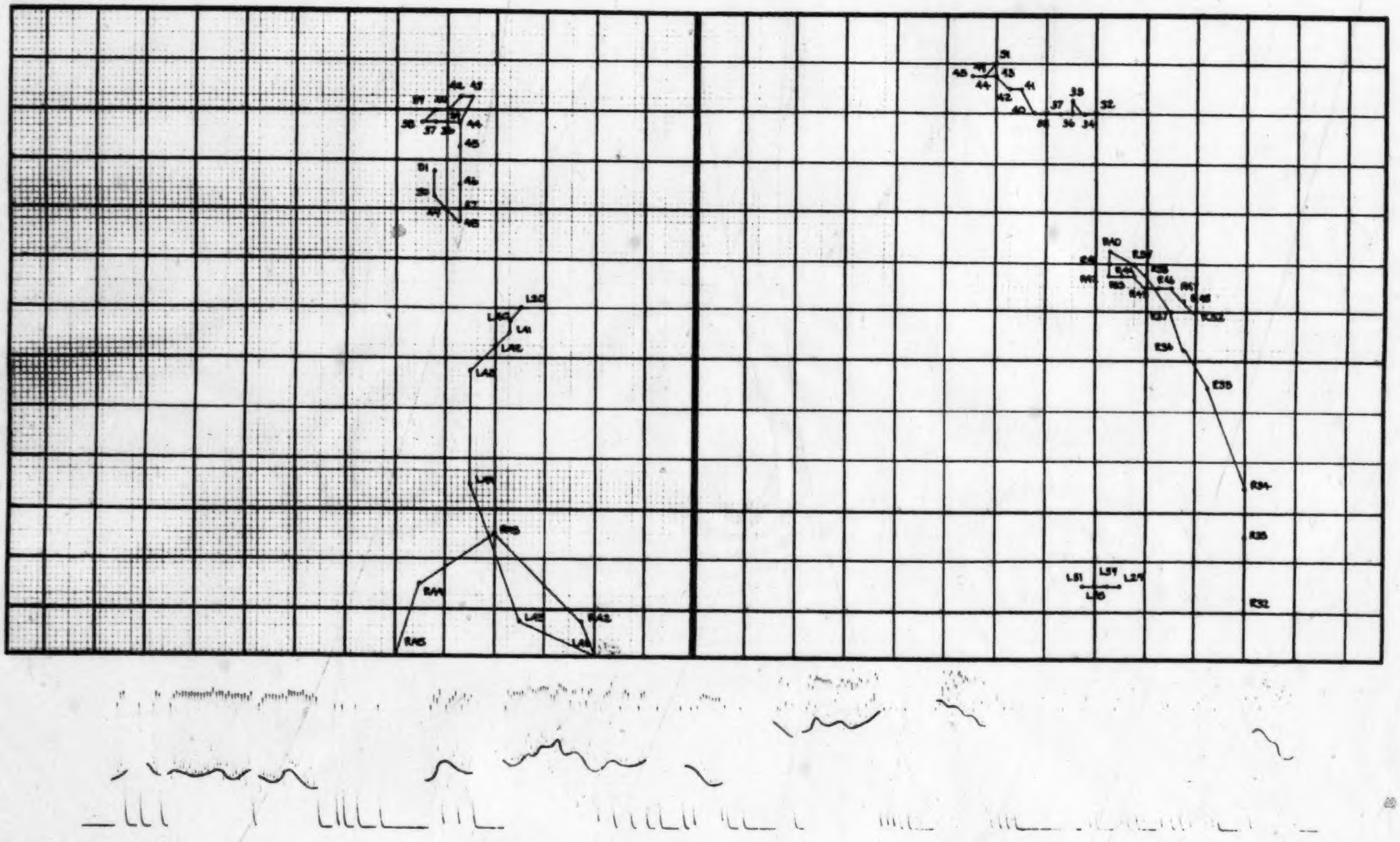
L grasps her cup just as C begins "I thought" (R29), but she does not begin to raise the cup toward her mouth until C has completed the utterance. It is, as the transcription indicates, in the silence between utterances that L instigates the series of head and hand movements necessary to obtain her sip of tea.

```
You little creeps! [laugh-----]
                   [laugh-----]
```

Example VI.3b

This segment is remarkable for the occurrence of coincidental and coetaneous sound and movement both within and between interactants. In particular the episode of mirth toward the end of the segment appears on the video screen as a terpsichorean production of sound and motion. There is some loss of motor control, which Mair argues is typical of "mirth," but in this brief interlude the rhythmic pulse of

Plate 6.12--Mingograph and Movement Traces: Example VI.3b



"°° E.C. -
f33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51
C: You little creeps! [laugh]
L: [laugh]

the interaction is largely maintained throughout. Interactants move in time with the pulse of laughter, and with each other. As L moves back and up, swooping around to the right (37 to 49), C moves forward and down (43 to 48). C almost literally doubles over with laughter. What is more, as the rather crude standard notation transcription of laughter indicates, they actually laugh together, coming in simultaneously on the beat (f44 to 46). There are, in addition small head movements effected by both interactants, separately and together, that, although too fine to be transcribed here, appear on the video screen to go with the sounds of their laughter.

You--The trajectory of the intonation contour is implied across the silence between utterances. The fundamental frequency picks up the trajectory at the point to which it would have got had there been no silence. It is a relatively flat intonation curve, falling slightly through the word/syllable until the last where there is a slight rise (f33-34 to 35). C herself is motionless here. L, however, moves to the right as the utterance is produced. Her head "bumps" up-right as the fundamental frequency rises (33 to 35). At the same time, her right hand continues its upward movement, in the process of bringing her cup to where she can drink from it (R33 to R35).

Little--The fundamental frequency falls with the first syllable, jumps up almost instantly, and falls in a similarly quick manner, through the same frequency range, with the second syllable (f35 to 37). C turns her head to the right, away from L, as she produces the utterance (35 to 37). She breaks eye contact momentarily, glancing down as she turns, but she re-establishes eye contact very quickly. There are no transcribable movements for C's hands.

L moves with the fundamental frequency of utterance here. Head movement parallels the motion of the fundamental frequency through the first syllable (35 to 36), and changes direction, shifting across-right, as the fundamental frequency jumps up to begin its fall through the second syllable (36 to 37). L's right hand continues to move up to the right (R35 to R37), bringing her cup ever closer to where she can take her drink.

Creeps!--Rising and falling, the fundamental frequency describes a long arch through "creeps" (f37 to 41). There is a good deal of "noise" in this utterance, voiceless consonants, and so forth, that is not recorded or recordable on the mingograph--such sounds have no fundamental frequency. Of especial note is the explosive /ps/ with which the utterance ends. It marks the beginning of the explosion of sound and motion within and between interactants that is here termed an episode of mirth.

C moves in time with the fundamental frequency of utterance. She turns her head back to the left in a movement that is coetaneous with utterance production (37 to 41). In addition, C's head "bumps" up-left (38 to 39), going with the intonation contour as the fundamental frequency peaks (f39-40). And, as the fundamental frequency falls, completing the utterance and delivering the state of play to momentary stability, C moves her left hand away from her neck and down toward her lap (L39 to L41). This motion appears to be preparation for and part of the laughter which follows immediately.

L moves in time with utterance here as well. As the fundamental frequency peaks, within 0.1 second of it peaking, L shifts her left hand to the right along the table top (L38 to L39), and achieves a drinkable position with the cup in her right hand (R39 to R40). The maximum frequency of the fundamental frequency occurs at the same time as L attains the maximum height (for this portion of the text, at any rate) of cup and right hand.

[Laugh--]--The fundamental frequency of these simultaneous bursts of laughter is clearly marked on the mingograph trace. The fundamental frequency is in a very high frequency range for both bursts. It dips, falling and rising through the first laugh(f44), and then is picked up, carried over the top, and brought down with the second (f46). The simultaneous terpsichorean movement of C and L through this segment begins precisely as C produces the plosive /ps/ with which she completes "creeps." As C finishes off this utterance she and L act. C moves forward and up, coming to face L (41 to 43). At the same time, L moves back and up (40 to 43). Like two ends of a stick, they move together, maintaining a more or less fixed distance from one another. During this few tenths of a second C begins to move both hands toward her lap. Her right hand comes up from the table, moving her body (R41 to R43), while her left hand continues to move down from her neck (L41 to L43). L moves her right hand and cup down, away from her head/mouth (R41 to R43).

C moves with the fundamental frequency of "ha." Her head veers down-right (43 to 44) as the fundamental frequency falls, and drops straight down (44 to 45) as the fundamental frequency changes direction and begins to move up. Her left hand effects a similar, coincident, change of direction, moving down (L43 to L44) and then down-left (L44 to L45). C's right hand executes a down-right shift, changing direction as the laugh begins (R43 to R45), although it continues to move toward her lap.

L also moves in time with utterance through this. Her head continues to move back, up, and around (41 to 50). The movement is rather poorly transcribed here, again because of the dimension problem. Both left and right hands move as the fundamental frequency changes direction i the first burst of laughter. L's right hand shifts to the left (R43 to R44) and then down (R44 to R45). Her left hand moves across-right (L44 to L45).

The transcription does not indicate similar coincidences of movement either within or between interactants on the second burst of laughter. Movement appears to continue on here. C and L "float" through trajectories already set in motion. Specifically, C's hands drop to her lap, slipping off-camera (R45 t R50; L45 to L50). She maintains the downward movement of her head, continuing through until two or three tenths of a second after "huh," is completed, when she begins to inhale in preparation for their next series of utterances (48). At that point her head reverses direction, moving up-right with the sound of her inhaling. L is still involved in the long swoop of her head (back up and around) that was initiated with /ps/. L's right hand continues to move horizontally, to the left (L45 to L48). Her left hand is motionless.

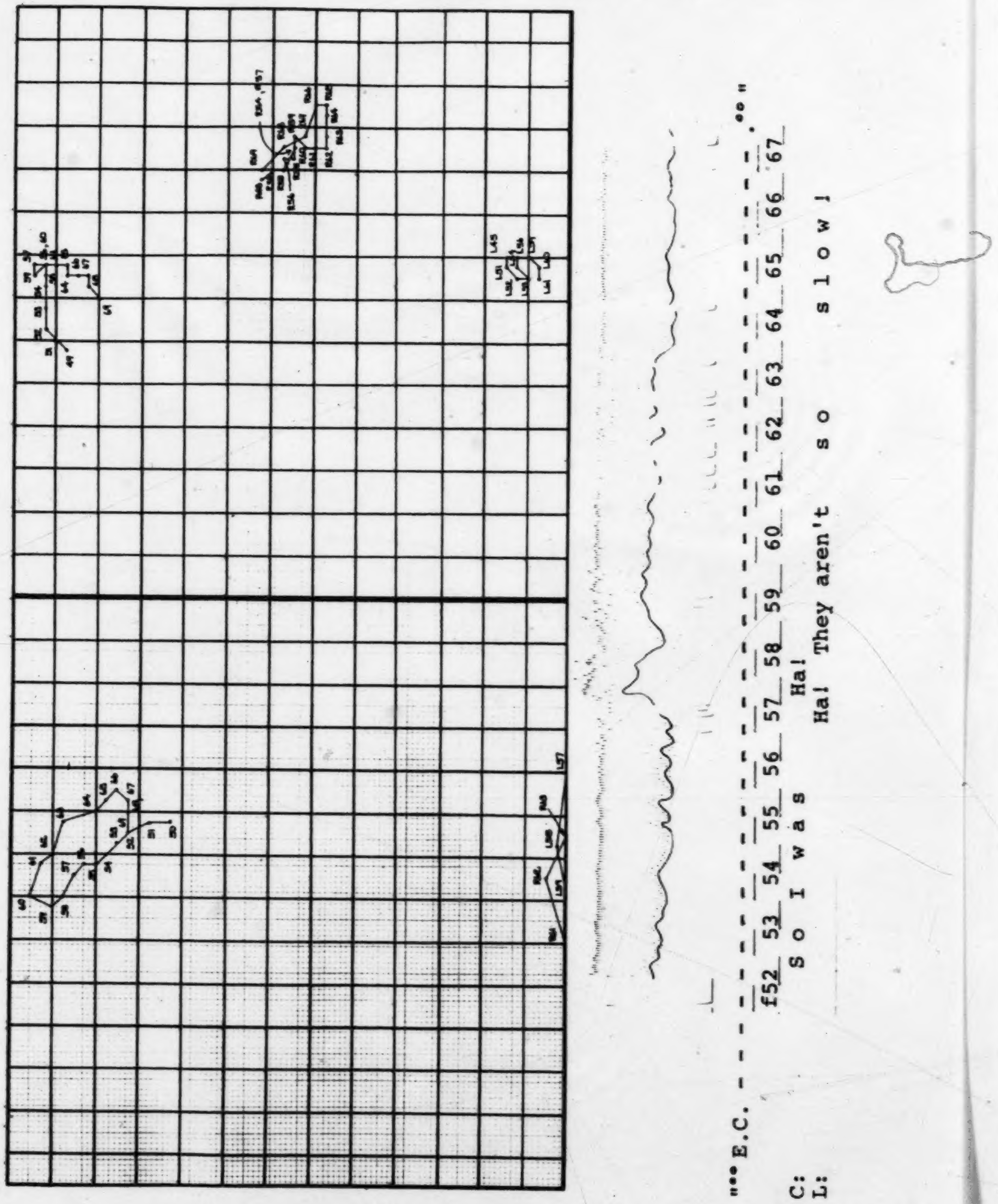
So I was . . . Ha!
Ha! . . .They aren't so slow

Example VI.3c

This segment carries on from the previous. The episode of mirth with which the preceding segment closes overlaps with and is continued on into this part of the text. Movements persist across the momentary silence that marks the arbitrary boundary which separates the two segments from each other. On tape, in real time, this and the previous segment appear as parts of one long multimodal outburst. Both interactants speak and move in time with the rhythm created by the fundamental frequency of utterance.

So I was--The intonation contour remains relatively flat through here (f51 to 57). Overall the fundamental frequency drops slightly through the utterance, until the

Plate 6.13--Mingograph and Movement Traces: Example VI.3c



end where there is a sharp rise. Specifically, the fundamental frequency dips slightly with each of the first two words, dips twice on "was," and rises quickly as the word is completed. The transcription shows that C's head moves up to the right during this utterance, continuing the motion introduced just after her second burst of laughter.

The transcription is misleading here in that it does not adequately indicate several head and body movements effected by C during production of the utterance. As C begins to move up-right, she turns her head to face L. That position is maintained and accentuated--C continues to turn her head to the left--through the utterance, Her body takes on what might be described as a "tough" stance: her shoulder is raised; head is turned; body is slightly tensed; she seems ready for action. On the video tape C's words here seem to be tossed out, as if in defense or defiance. In addition, there are small movements of C's head--a sort of nodding--too fine for transcription that seem to go with the utterance to the level of the syllable.[13] Finally, it is apparent on the video tape that most of the large movement of C's head is effected by or is part of the movement of her torso. C pivots from the waist throughout this segment, swinging her upper body through precisely controlled arcs, in time with utterance.

---

[13] This bit of description in unabashedly subjective and impressionistic. It takes off from the data and is offered here only to add depth to and contextualize the data presented in the mingograph traces and transcription of head and hand movements.

L also pivots from her waist through this segment. She is in the process of describing a large "loop" with her upper body. The motion, begun a second ago with "creeps," proceeds clockwise up to her right, back around and down-right. The transcription here indicates only head movements up-left (50 to 52) and left (52 to 56). These are the "back around" portion of the "loop." L's right hand still holds her cup, almost in a state of suspension between drinking and not drinking. Right hand motion generally parallels the motion of the fundamental frequency. The right hand/cup combination moves down-left with the overall drop of the intonation contour (R51 to R55), and ticks up as the fundamental frequency rises at the end of "was" (R55 to R56). Although L's left hand remains on the table top and moves only infrequently, movement seems to go with utterance and upper body motion. That is, her left hand shifts down-right with the falling fundamental of "so I" (L51 to L53), and moves up-left as her body swings back around (L53 to L57).

[Laughter---]--These two laughs are virtually coincidental; the occur with 0.1 second of one another. As laughter, they seem to carry on from the previous episode of mirth. (At the time of their occurrence, the smiles are still in place on the faces of both interactants.) The fundamental frequency trace clearly marks the outburst of laughter (f57-58) but it is not possible to completely distinguish the two outbursts on the mingograph record.

Attribution of a particular trace to a particular interactant's laugh must appeal to audio and video recordings. In real time L's laugh is heard to occur just before C's. L's laugh would then appear to be indicated by the first arch of the high frequency fundamental. It rises and falls sharply immediately after C completes "was" (f57). C's laugh is then indicated by the second arch of the high frequency fundamental. It too rises and falls dramatically, just preceding (as the ear would agree from the audio tape) L's next utterance (f57-58).

L moves in time with the pulse of utterance. The transcription indicates an upward motion of her head in the 0.1-second before her laugh (56 to 57). Once again, however, the transcription is inadequate to the reality. On the video tape the movement looks more like a "hiccup" of motion. L brings her chin in to her chest and raises her shoulders and upper body as she raises her head. It is a single complex of motion, occurring almost as she laughs. L's right hand moves down (R56 to R57) with her laugh, and then proceeds left, as C laughs (R57 to R58).

On the audio tape C's laugh sounds less emphatic, less strong than L's. This admittedly subjective impression goes well with what is seen on the video tape, and with transcribed movement through this portion of the text. C continues to move head and torso up to the right (56 to 58). There are no quick movements or changes in direction. C's left hand parallels the movement of the fundamental frequency, rising and falling with it (L56 to L58). Left hand slips off camera, to come to rest on the table top, just after, with 0.1 second after, L has commenced "They aren't so slow."

They aren't so slow--C has her mouth open and is leaning forward, toward L, as L produces this utterance. C looks as thought she is about to speak. L, however, anticipates the beat and enters with her sentence before C can get back into the conversation. From that point of entry L's utterance continues, as the mingograph traces show, in time with the established rhythm and melody of the interaction. And both interactants continue to move in time with the fundamental frequency of utterance. L appears to almost "float" gently down from the "hiccup" of her laugh as she produces "They aren't so slow." Head and torso movement parallels, going with the general direction of movement of the fundamental frequency. Except for this steady downward movement of head, hand, and upper body there is a remarkable lack of transcribable movement on L's part through the utterance.

C, on the other hand, is more active. She completes a movement sequence that began with "so I was." As this utterance begins C swings head and torso clockwise down, back, and away from L. At the same time she turns her head to the right and looks down, lowering her head, to break eye contact. She looks up very quickly, within 0.1 second or so, and re-establishes eye contact. C completes her loop as L's utterance ends. She rises toward L, facing her. On the screen C appears to be in the process of setting up for her own utterance, which follows immediately. C leans forward slightly, wets her lips, and opens her mouth. She is ready and waiting.

They--The fundamental frequency carries on from the very high frequencies of the laughter, but quickly drops through the first part of the word. The mingograph, in accord with what the ear hears, shows that the fundamental frequency bottoms out in mid-word, is turned around, and moves up slightly through the latter portion of "they" (f58 to 59). L moves coincidentally with the initiation of utterance. Three movements occur simultaneously: her head goes to the right (58 to 59); right hand moves left a bit (R57 to R59); and left hand shifts down, moving away from her body, along the table top (L58 to L59).

Through this 0.2 second C continues the loop of head and torso, moving back and to the right (58 to 59). At the same time, in a complex of simultaneous movement that is too fine for proper transcription here, C turns her head to the right, tilts her head forward, and looks down, breaking eye contact.

Aren't--The fundamental frequency continues the upward motion which originates in the latter part of "they," is carried over the top early in the utterance, and falls a bit through the rest (f59 to 61). The frequencies of the fundamental through this are at higher than normal levels.

They approximate those of the more emphatic "creeps."

It is here, with "aren't," that L begins to float down from the height of her "hiccup." She moves down with the falling fundamental frequency. Specifically, head movement proceeds down-left (59 to 61). Her left hand moves down-right along the table top (L59 to L61), and comes to rest palm down. Her right hand goes very closely with the intonation contour, shifting down-right to down (R59 to R61) precisely as the fundamental frequency changes direction.

C, meanwhile, is into the back around portion of her loop. The transcription marks her head movement as down-left (59 to 61). Again, though, transcription of a single spot on the face does not adequately indicate the complexity of C's head movement. From the video tape it is apparent that C turns her head back to the left and, lifting her eyes, re-establishes eye contact in this 0.2 second. C is moving toward L once more.

So--the fundamental frequency arches through "so," peaking roughly in mid-word (f61 to 64). It is a relatively long word/syllable, during which the fundamental frequency comes down from the higher frequencies of previous parts of the utterance to a frequency range more comparable with the majority of the discourse through this fragment of text. In addition, it should be noted that there are a number of zero-line indications on the mingograph trace that do not correspond to the sound the ear hears on the audio tape through this portion of text. They may be the record of voiceless consonants, /t/ or /s/; or they may indicate a weak audio signal.

L continues to float down through "so." Head continues down, veering to the right, changing direction with the fundamental frequency (61 to 64). Her right hand moves down and to the left, also changing direction as the fundamental frequency peaks and tumbles over (R61 to R64).

C has completed the back around portion of her loop here. The transcription indicates that head movement continues down-left (61 to 64). C is in eye contact, facing L, and moving toward her. It is here, at this point in time, that C wets her lips. At the same time she shifts her right hand across the table toward L. Right hand movement describes an arc (R60 to R64) that rises and falls with the fundamental frequency. Precisely as L rounds off "so," C's right hand comes back to rest on the table top.

Slow--The fundamental frequency contour is generally flat through "slow" (f64 to 67). There is, however, some variation evident on the mingograph trace. There are two small ripples or arches in mid-word; and the fundamental frequency rises as the utterance ends. The brief silence marked by the mingograph at the completion of "slow" corresponds to, corroborating, what is heard on the audio tape.

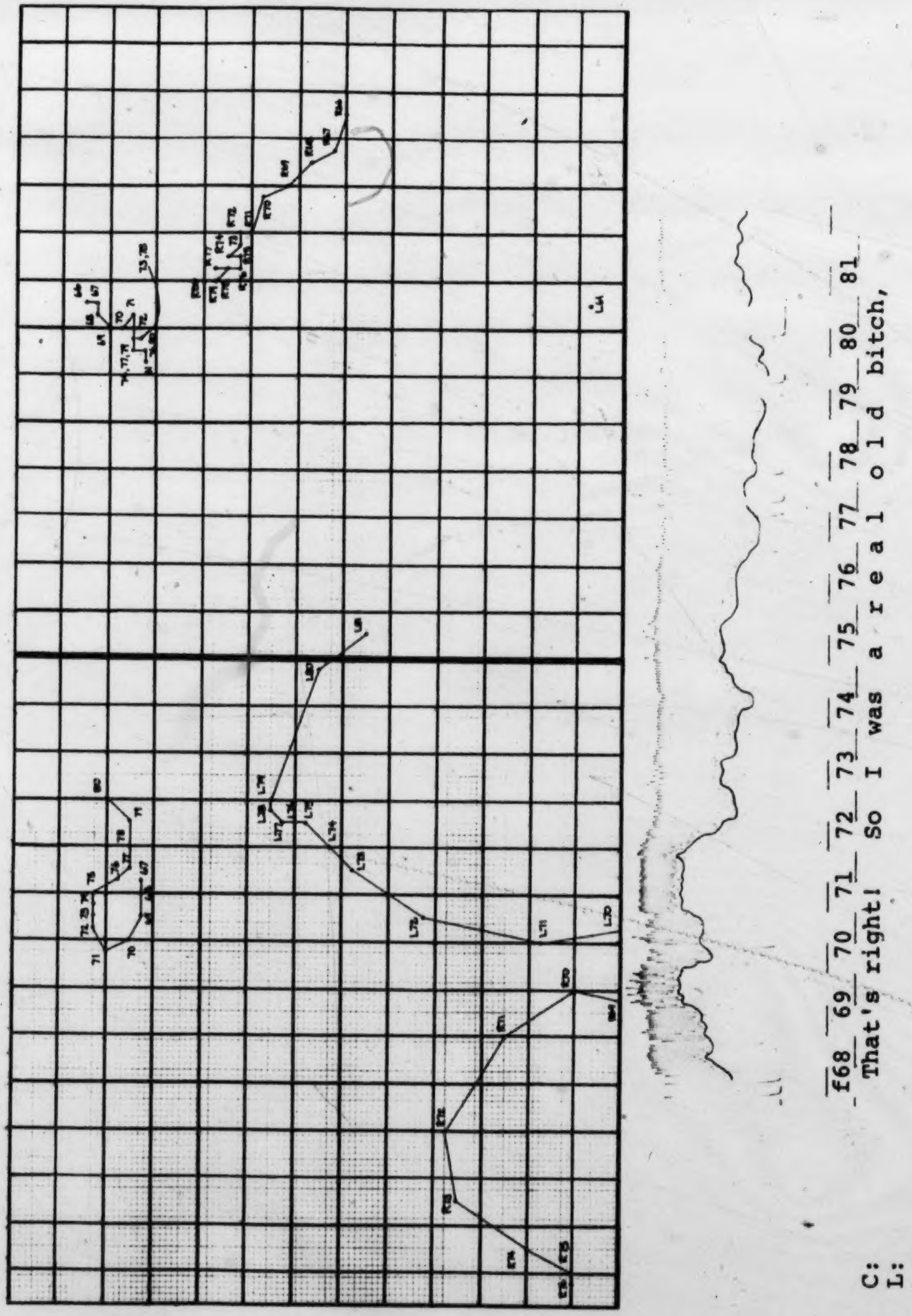
L completes her float down during "slow" (64 to 66). Just as the utterance ends, L begins the series of movements that are at last to result in her sip of tea. These movements are indicated in the transcription as the down movement of her head (66 to 67). Her right hand evidences a similar change in direction 0.1 second into "slow" (R65 to R67). Anticipating head movement by 0.1 second, this marks the initiation of movement of right hand and cup toward L's mouth. Right hand moves up and to the right here.

C also completes a movement sequence during "slow." She finishes the loop described by head and torso movement. As the transcription indicates she moves down-left through the first 0.2 second of the utterance (64 to 66) and, precisely as the fundamental frequency begins to rise (at the end of "slow"), C initiates a move away from L. In the transcription this is marked as a shift down-right (66 to 67). C maintains the direction of her regard through this short segment. Eye contact is maintained only briefly into the next segment, however. L breaks eye contact as she looks down to her cup and sips her tea. Finally, in this segment C's hands remain off-camera, resting motionless on the table top.

That's right! So I was a real old bitch.

Example VI.3d

This segment is much quieter and more reserved than the previous two. The mirth in evidence through the earlier segments is finished. L at last takes time to sip her tea. C speaks alone. It is proposed that L "gives over" control, Plate 6.14--Mingograph and Movement Traces: Example VI.3d



allowing C to complete her story of the students at the junior high school.

C and L are in eye contact as the segment begins. L breaks eye contact first; she glances down to her cup as C produces "that's" (67 to 68), and, a moment later, when she looks up, C is no longer looking toward her. C looks down and away from L as she produces "right" (69 to 71). For at least one full second neither C nor L is able to re-establish eye contact. When one is looking, the other is not. It is only in the segment which follows this one, as C begins "and they behaved themselves..." that she looks up to catch L's eye and eye contact is re-established.

That's right--The fundamental frequency rises sharply and dramatically with "that's" (f67 to 69). It is a high-pitched and emphatic utterance. The fundamental frequency arches up and over through "right." As "right" is completed, the fundamental frequency tumbles over and down toward the more normal range of C's vocal production (f69 to 71).

C moves emphatically with utterance here. She turns her head quickly and sharply to the right as she produces the utterance (67 to 71). The upward movement indicated in the transcription is the result of back-up movement of her upper body which she produces as she pivots in her chair (69 to 71). Although also executed in time with utterance, this movement is not properly a "head movement" at all. C's hand movement follows the direction of motion of the intonation contour through time. She raises both hands from her lap precisely as the fundamental frequency peaks in "right" (R69 to R71; L70 to L71).

L also moves with the fundamental frequency of utterance here, although less dramatically. L shifts her head to the right slightly as the fundamental frequency rises sharply through "that's" (67 to 68), and slips down to her cup with the falling fundamental frequency of "right" (68 to 71). Her right hand moves up and right, bringing cup to mouth (R67 to R71). Left hand remains motionless on the table top.

So I was--The fundamental frequency rises quickly with "so" (f71 to 72), arching through very high frequencies similar to those of the previous utterance, and then tumbles down to the more normal frequency range for C's utterances. "I was" has a relatively flat intonation contour (f72 to 74). There is, nevertheless, a small rise with "I" and a slight drop to round off "was."

C abruptly reverses the direction of her head movement, from across-left to across-right, as the fundamental frequency peaks in "so" (71 to 72). The smooth continuous movement of head and torso up and to the left which follows goes well with the relatively flat intonation contour of the balance of the utterance (72 to 74). Hand movement goes more particularly with the motion of the fundamental frequency. C's right hand, already rising from her lap, peaks as the fundamental peaks in "so," and turns with it to fall down toward the table (R71 to R74). Her left hand continues to move up-left toward her chin (the position in which it began this fragment, L71 to L74).

L's movements also go with the relatively flat intonation contour here. That is, movement is relatively flat and unremarkable. There are no large or dramatic gestures. L moves her head slightly to the right (71 to 73), and brings her right hand up-right a bit (R71 to R74). Here as the utterance ends, L finally puts cup to lips and presumably drinks.

A real old bitch--The fundamental frequency contour arches prominently through "a real" (f74 to 77), rising quickly with "a" and dropping off with the last part of "real." In this utterance the fundamental frequency drops to its lowest frequencies in the fragment; only "I thought," near the beginning of the fragment, has a fundamental frequency in a comparable frequency range. During production of this "a real" C continues the several movements undertaken in previous segments of the text. She completes the "back" portion of her head/upper body loop: head movement goes left with "a" (74 to 75) and swerves down-left, toward L, as the fundamental frequency peaks in "real" (75 to 77). C's right hand falls to the table, coming to rest as the fundamental frequency tumbles over and moves down through "real," as well (R74 to R76). Her left hand continues to describe a long arc, moving up-left toward her chin (L74 to L77). The small movements transcribed for L here indicate "drinking" motions (74 to 77). The solitary downward motion of her head occurs precisely as the fundamental frequency comes over the top and begins to fall through "real" (75 to 76).

On the mingograph trace and on the audio tape, "real" flows directly, without pause, into "old." Both are characterised by extremely low fundamental frequencies. the mingograph documents a noticeable arch of the intonation contour through "old" (f77 to 79-80). With "bitch" the fundamental frequency climbs to a more typical range for C. The fundamental frequency is sustained there, maintained through to the silence that marks completion of utterance (f79-80 to 82).

C moves with the fundamental frequency of utterance here, again. As noted above, C has completed the long series of looping movements. With "old" C moves to the left, toward L (77 to 79). It is a short, brief motion. Almost immediately she veers back up and to the right, moving up with the rising fundamental frequency of "bitch" (79 to 82). C reaches the apogee of the arc described by her head movement (82) precisely as the utterance ends, in silence. In that same 0.1 second, C turns her head in L's direction and re-establishes eye contact. Although C's right hand remains motionless on the table, resting beside her coffee cup, her left hand continues up toward her chin through "old" (L77 to L79), and then reverses direction, proceeding down-left and off camera with "bitch" (L79 to L82).

L sips her tea through these utterances, executing, as the transcription indicates, only small movements with her head and right hand (77 to 82). Perhaps most notable, and this on the video tape, is L's head movement: she moves to her right, away from C, as C produces "bitch."

And they behaved themselves--C speaks alone, completing the story of her experiences as "sub" at that vocational junior high school. As noted above, it is thought that L has given over control to C. L has apparently chosen to be "listener" for the balance of the story. She does not opt out, however; she moves in time with the fundamental frequency of utterance right through to the end of the story and the fragment. She continues to take an active part in the evolving system.

And--The trajectory of the intonation contour carries on across the brief silence which separates the segments and is taken up at the point where it would have got to had there been no silence. With "and," then, the fundamental frequency drops off rapidly, completing the S shape of the contour (f82-83 to 83-84). C moves with utterance. Her head turns down a bit as the fundamental frequency drops (83 to 84). Both C's hands are off camera at this point. Her right hand rests on the table top; her left hand is still dropping down it.

L is just finishing her sip of tea here. As the utterance is produced, she moves her head and right hand simultaneously. Head shifts up (83 to 84) and right hand moves down (R83 to R84), moving very particularly with the falling fundamental frequency of utterance.

They--The fundamental rises fairly quickly through "they" (f83-84 to 84-85). This is a multiple pitch production, almost a two-syllable utterance. The fundamental frequency arches through the first portion of the word, falling a bit, though not back down to the frequency from which it began (f83-84), and then turns to rise again to the peak of a second arch as the word is completed (f84-85). In this 0.1 second there are but two transcribable movements. L's head moves up very slightly (84 to 85), going with the rising fundamental; and C brings her left hand down to the table top, clicking her ring on the hard surface there (L84-85).

Behaved--With the first syllable of the utterance the fundamental frequency drops, almost completing the arch begun with "they" (f84 to 86). The fundamental frequency continues to drop slightly through the first bit of the second syllable, moving to relatively low frequencies, and then rises, arching upward as the word ends (f86 to 87-88).

There is very little movement during this utterance. L is in the process of putting her cup back down on the table. Her right hand moves down-left (R87 to R88). L's head moves down a bit (85 to 86) and left (87 to 88). C moves her head to the right slightly (85 to 87) and nods (86 to 88). Both C and L change the direction of the respective head movements coincidentally, within 0.1 second of the change of the direction of the fundamental frequency. C and L are moving away from each other here; it is a direction of movement that continues to the end of the fragment.

Themselves--The fundamental frequency rises slightly through "themselves," arching through the utterance (f87-88 to 90). The word ends on the high point of the S-shaped curve described by the intonation contour. The instability attributed by the model to that state of affairs goes well with the expression on C's face and the position of her head and torso--she is leaning forward, toward L, off balance.

C continues to nod through this utterance (87 to 91). It is a movement sequence she maintains through to the end of the fragment.

L also carries on with movements already in progress. Her head movement is to the left and up, going with the fundamental frequency (87 to 91). Her right hand continues on its journey, escorting her cup back to the table top (R87 to R91).

And they did some work--There is no record of fundamental frequency for this segment of text. C is far from the microphone; she has moved her body to the left and turns her head to the left, as well. What is more, she produces the utterance at very low volume. The recorded audio signal is consequently too weak; the trans pitchmeter and mingograph can not extract fundamental frequency from it. Perhaps fortunately, there is very little movement by either interactant during this time. C continues to nod as she speaks, finishing up her story. L still moves to her left, away from C. There are a couple of movements of note, however. Precisely as C completes her utterance, L touches her cup down on the table top (R100). And precisely as C completes her utterance, she freezes for an instant. She is in direct eye contact with L. L also pauses momentarily. It is one of Mair's "timeless moments." Each realises that the other has realised, and so forth, to infinite regress. The moment, lasting perhaps 0.1 second, cannot be said to stop the text; it is more properly a hesitation. L breaks it off, redirecting the text into a new topic with her question: "What were you subbing?".

6.4.3 Sound and Movement Clusters--Interpretation

In this brief fragment of conversation there is a multitude of instances of coincident and coetaneous sound and movement both within and between interactants. Such instances are argued to support the notion of synchrony in communicative interaction. There are shared movements and timings, complex interweaving of melody, and coming in together after silence. The episode of mirth toward the middle of the fragment is especially obvious and striking, and of particular note for this study. C and L seem to "explode" all over the video screeen and audio tapes at this point. Yet upon closer examination it is clear that they may be more properly described as partners in a complex, intricately timed dance, moving in time with the pulse of

Plate 6.15--Mingograph and Movement Traces: Example VI.3e

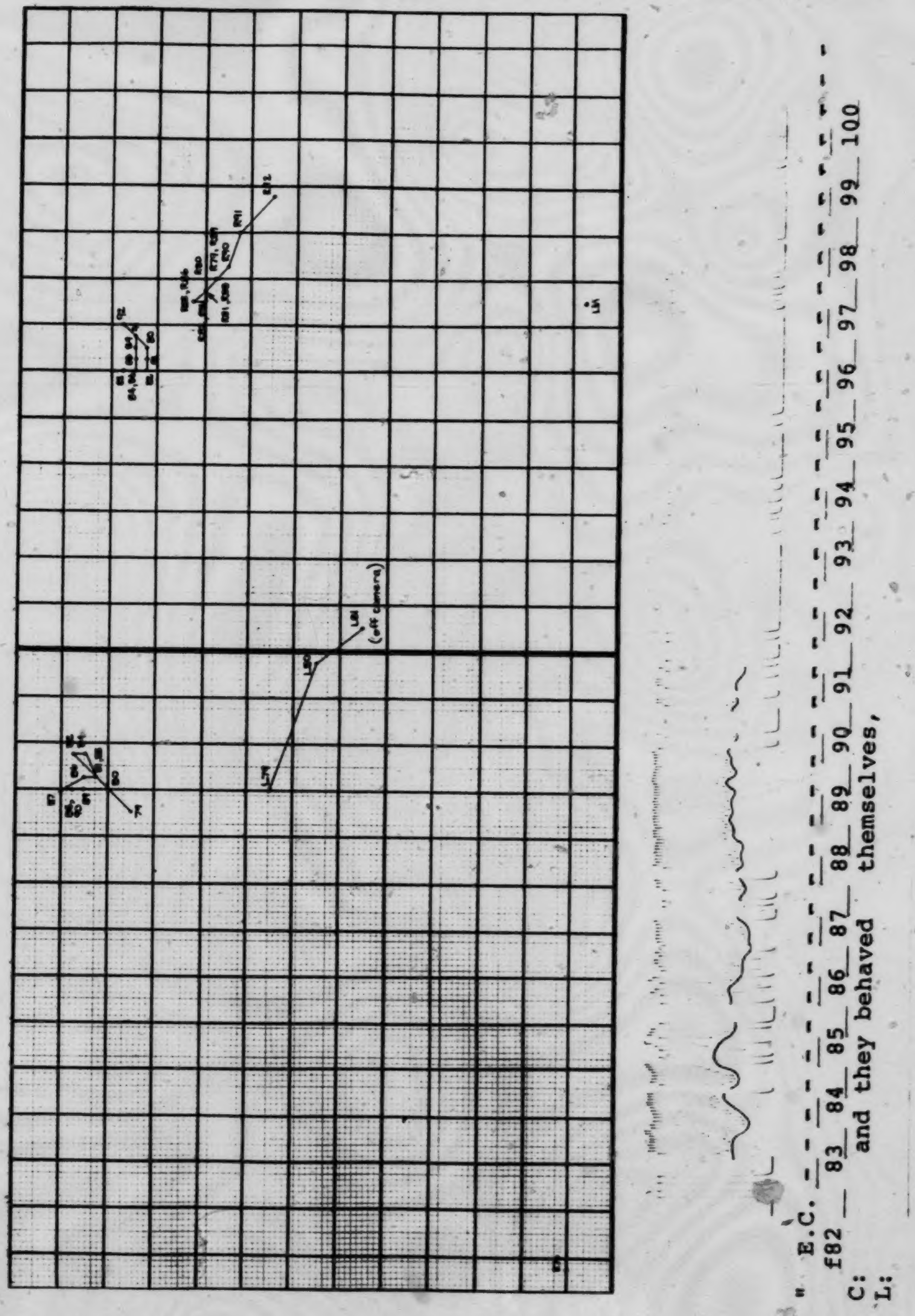
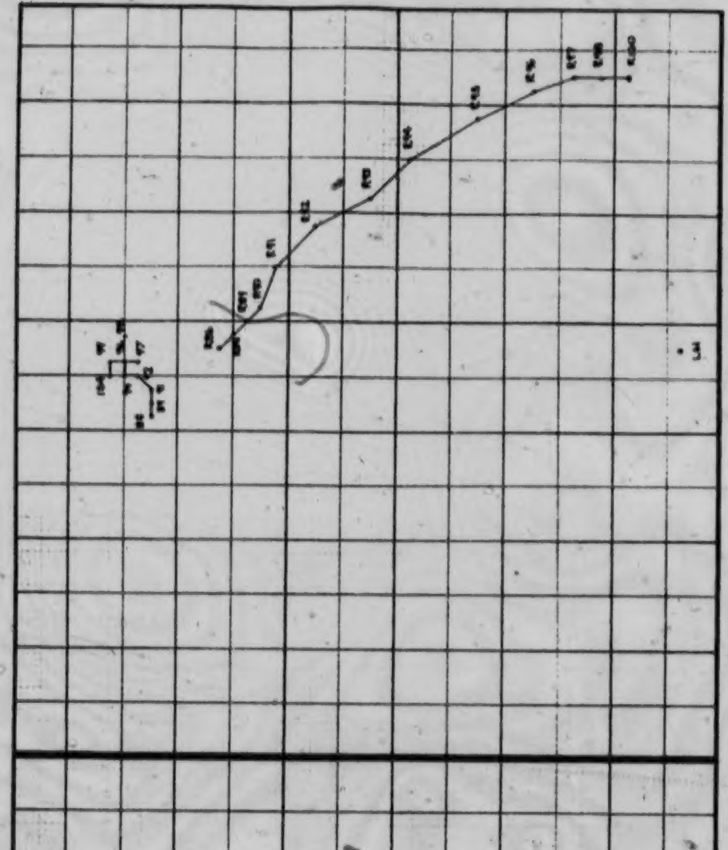


Plate 6.16--Mingograph and Movement Traces: Example VI.3f



C: And they did some work.

the words and laughter.

As before, the description shall be interpreted from the point of view of Mair's model. It will be argued that the configurations of sound and motion singled out for consideration in fact constitute evidence for the existence of intra- and interpersonal synchrony as an identifiable, observable, and definable aspect of human interaction. Following the procedure established in previous examples, I will, where possible, deal with each segment separately, focusing specifically on those clusters of sound and movement held to exemplify the notion(s) of intra- and interpersonal synchrony in human interaction.

```
          You know, and after that I thought
    [laugh------------]
```

Example VI.3a

<u>[Laugh]</u>--L moves with the pulse of her laughter through this portion of the text, shifting both head and hand precisely as each pulse of the laugh occurs. The fine integration of sound and movement within an interactant is taken here as evidence of intrapersonal synchrony. The events in each modality seem to be part of, and dependent on, events in the other; both appear as integral aspects of the state of play being delivered with their occurrence.

C also moves in time with the pulse of L's laughter through this portion of the text. Head and hand movements occur as the laughter is produced. This integration of sound and motion between interactants is thought to provide support for the concept of interpersonal synchrony. Both interactants appear to be moving through time together, taking the timing of their own cognitive processes from the shared rhythm of utterance. The simultaneous break of eye contact at the end of this segment is also argued to constitute interpersonal synchrony. The very small time frame and the detailed precision of the movement obviate arguments for a visual or aural cueing system. It is further proposed that such coinstantaneous action may be most adequately and accurately accounted for as part of a "shared awareness": interactants are immersed in the interaction, nodes of single system in evolution.

You know--As C produces "you know" she is in the process of turning her head to the right. She shifts her head up, without interrupting horizontal motion, precisely as the fundamental frequency peaks. This simultaneous shift in sound and motion is thought to support the notion of intrapersonal synchrony.

L also moves with the fundamental frequency of utterance. Her head movement parallels the movement of the fundamental frequency, slipping down with the fundamental frequency of "know" and across-left as the fundamental frequency rises. Her right hand, moving across the table top to her cup, changes direction coincidentally with the changes in direction of the fundamental frequency. Such coinstantaneous manifestation of sound and motion between interactants is interpreted as evidence of interpersonal synchrony in action.

Furthermore, the fact that this utterance by C comes in within 0.1 second, almost literally on top of L's second burst of laughter, is thought to be another example of interpersonal synchrony in interaction. It all happens far too quickly and with far too much precision to be accidental, or to be the product of a feedback system. Again, C and L are thought to be immersed in the shared text of interaction, aware of and taking the timing of their cognitive processes from the rhythms impinging on their sense, the shared text itself. They are working through a shared reality together, in time.

And after that--As C produces this utterance she is in the process of re-turning her head to where she is facing L once again. C's head moves back-left with the utterance, following the long arc described by the fundamental frequency. The sound and movement appear on the video tape as one, and are argued to be products of a single central patterning program in the brain, markers of intrapersonal synchrony in action.

During these same 0.3 second, L swings her head to the left--she too once more faces her partner--and shifts her left hand down-left, paralleling the movement of the the falling fundamental frequency of "that." As they are coinstantaneously completing their respective head movements, C and L simultaneously seek and re-establish eye contact. Both the coincidence and the fine integration of sound and movement between interactants are thought to support the argument for the existence of interpersonal synchrony. As with all such examples, time frame, level and detail of coordination between modalities and interactants demand explanation in terms other than those of a stimulus-response model.

I thought--C is motionless here. There is, then, no evidence of intrapersonal synchrony in this bit of text. L, however, does move in time with C's utterance. L's head goes with, imitating, the rise and fall of the fundamental frequency. Her movement is here interpreted as interpersonal synchrony.

You little creeps! [laugh-----]
[laugh-----]

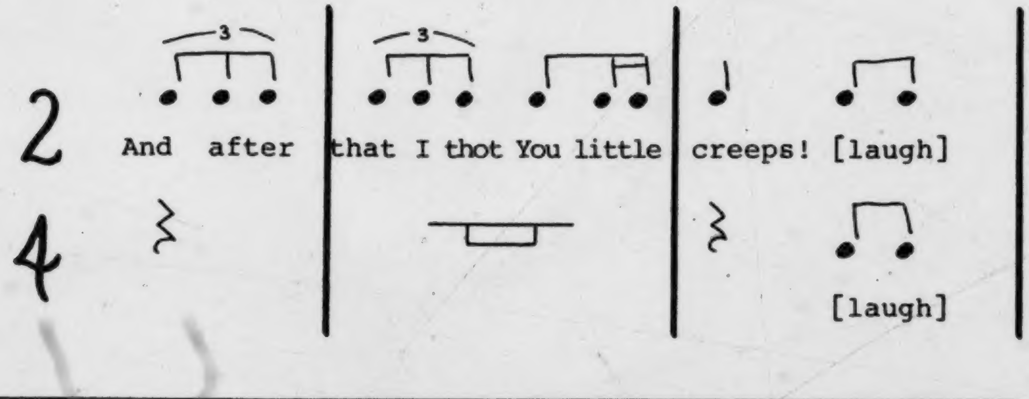
Example VI.3b

As noted in the description, this segment is a remarkable one, particularly for the coincident, coetaneous "explosions" of sound and movement that populate the episode of mirth with which the segment ends. The fine integration and close coordination of aural and visual modes within and between interactants here is argued to support the notion of "synchrony" in all forms. It is apparent from the data that C and L move through time in time--that is, with each other and with the pulse of utterance--during this segment. C bends forward and down, doubling over with her laughter, precisely as L loops back and around. Their laughter is coincidental and coetaneous. C and L come in together. They laugh simultaneously on the beat. In fact, the sounds of their laughter seem to be the "music" to which they dance out the segment. Figure 6.7 is the Western musical notation that describes the rhythm of the segment. In so far as the data are a record of events occurring between interactants they are evidence of interpersonal synchrony. In so far as they mark events within interactants they are evidence of intrapersonal synchrony. There are, in addition, very small head movements by both interactants, too fine for transcription on the scale used here, that seem to go with the fundamental frequency of utterance as well. It is proposed that such movement is the product of a single central patterning program in the brain and as such may be seen as intrapersonal synchrony in interaction.

You--Once again, C is motionless here while L moves in time her (C's) utterance. L moves her head up and to the right as the fundamental frequency rises, paralleling its movement. At the same time, she brings her right hand up, cup in hand. This coincidence of sound and motion between interactants is thought to constitute evidence of interpersonal synchrony.

Little--C moves with her utterance here, swinging her head to the right in a motion that is coincidental and coetaneous with utterance. The precise integration of aural and visual modes and the very small time frame involved in this transmodal production suggest that such simultaneous

Figure 6.7: Musical Notation for "...little creeps"



manifestation of sound and motion within an interactant might be best accounted for as intrapersonal synchrony. L continues to move with the utterance here as well, providing further examples of what are interpreted as interpersonal synchrony. L's head movement is down through the first syllable, paralleling the motion of the fundamental frequency. Head and fundamental frequency change directions coinstantaneously with initiation of the second syllable of the utterance.

Creeps--That C moves with the fundamental frequency of her utterance, producing sound and movement together in time, is taken as evidence of intrapersonal synchrony in action. Specifically, C turns her head back to the left through "creeps" and without interrupting horizontal movement raises her head slightly, precisely as the fundamental frequency peaks. In addition, as the fundamental frequency turns and falls to complete the utterance, C drops her left hand to her lap.

L's movements through this portion of the text also occur with C's utterance, and are taken as evidence of interpersonal synchrony in interaction. L achieves maximum height of the cup-right hand constellation and shifts her left hand along the table top just as the fundamental frequency peaks.

[Laugh]--There are two simultaneous bursts of laughter in this sequence. They are accompanied by and go with a series of precisely timed, coincidental and coetaneous movements on the part of both interactants. On the video screen it appears as a "dance" of mirth. As such it is thought to offer support for the existence of synchrony in interaction. The simultaneity is in fact what motivated the selection of this fragment for description. Because the transcribable sounds of laughter occur simultaneously on the audio and video tapes and the mingograph traces, it is not possible to attribute sound to a particular interactant. It is consequently rather difficult to define intra-, as opposed to interpersonal synchrony in the data. The more general term, synchrony, must suffice and will be employed in discussing this portion of the episode.

The terpsichorean coincidence of sound and movement of interest actually begins with the explosive /ps/ with which C finishes "creeps." C and L begin quite obviously to move together as the "explosion" of sound is released. C moves forward and up as L moves back and up. They move in parallel fashion, each performing coincidentally the mirror image of the other's movement. This is argued to be synchrony in action. There is also what is taken to by synchronous action in the movement of the extremities of both interactants. That is, C executes a series of hand movements, bringing them toward her lap, while L moves her right hand and cup down and away from her mouth.

The first burst of laughter continues with the "dance." Both interactants move in time with utterance and with each other. The detailed coordination of sound and motion is put forward as evidence of synchrony in interaction. Everything happens at once; yet the eruption of sound and movement in the seemingly chaotic disarray of "now" is neither chaotic nor disarrayed, but rather, precisely timed and coordinated in both aural and visual modes, within and between interactants. C's head moves down-right with the falling fundamental frequency, and changes direction, dropping straight down, just as the fundamental bottoms out and starts to climb. C's left hand effects a similar, coincident change of direction. Her right hand swerves with the onset of laughter. L's left and right hands also change direction. Her left hand shifts down, right hand veers across-right, as the fundamental frequency reaches its minimum amplitude and turns to rise. L's head and torso persist in the larger looping motion begun with the initial explosion of sound. In all, it is a synchronous symphony of sound and motion, taking place extremely quickly and with exact control. Feedback loops and/or stimulus-response models cannot account for temporal and physical precision (aural and visual) on this scale.

Although the "dance" continues, there is no similar eruption of coincident, coetaneous manifestation of sound and movement within and between interactants with the second burst of laughter. Movement seems to more or less carry on, pursuing trajectories initiated with the first explosion of mirth. C proceeds with the movement of her head and torso, bending forward and down. Her hands continue to drop to her lap, slipping off camera. L is still completing the loop (back, up, and around) of her head and torso. Her right hand sustains its horizontal motion. Both interactants are in a state of "continuance." The continuation itself constitutes a kind of synchrony; that is, the movements of C and L continue to occur in time with one another and with the fundamental frequency of utterance. They are still dancing. The lack of any "eruption" of sound and motion through this portion of the text may be, at least in part, the product of a partial lack of motor control, purported to accompany mirth in interaction. At any rate the continuity of movement through sound evidenced here clearly demonstrates that it would be futile to search for invariant relationships between movement and utterance. They are always realised in context.

So I was . . . Ha!
          Ha! . . .They aren't so slow

Example VI.3c

As in all examples, there is overlap across, and continuation between, segments in the fragment. Here in particular movement and mirth persist across arbitrary segment boundaries. The state of play being delivered with their occurrence takes place over the boundaries established for microanalysis. This segment itself is notable for the smooth sequencing of repartee between interactants. L might be said to "interrupt" C's story with a sidetrack of sorts, but her utterance comes in on time and carries on the established character of the fundamental frequency of utterance. The beat goes on. This kind of behaviour between interactants is thought to demonstrate interpersonal synchrony in interaction. Both parties move and speak in time with each other, with the rhythmic divide of time effected by the falling S shapes of the fundamental frequency of utterances produced by both interactants.

So I was--C moves simultaneously with production of her utterance. Head and upper body movements are "onomatopoeic"; that is, movement describes a long, smooth arc up to the right that is similar to the arc described by the fundamental frequency contour. Such coinstantaneous production of sound and motion within an interactant is thought to be most adequately and accurately accounted for as the product of a single central patterning program in the brain, and is here interpreted as intrapersonal synchrony.

There are also smaller movements of C's head that are produced in concert with utterance. C turns her head to face L and nods as she speaks, "tossing" her words at L. Although not transcribed, and not transcribable with the present scale of measure, the fact of these movements on the video tape is argued to provide additional and supplementary support for the notion of intrapersonal synchrony.

L's head and upper body move through part of a similar looping arc, though in the opposite direction, during this utterance. Her motion, then, is also "onomatopoeic," describing a long, fluid arc along with the fundamental frequency. L's right hand parallels the fundamental frequency contour as well, moving down with it through the utterance to the last 0.1 second where both "jump" a bit. L's left hand follows the motion of the fundamental frequency also until this hand movement is "interrupted" by the swing of her torso and must shift to go with it. Such perfect congruence of sound and motion between interactants is thought conclusively to demonstrate interpersonal synchrony in action.

On the bursts of laughter, C and L come in together, within 0.1 second of each other, for this brief episode of mirth. The two laughs occur almost literally the one on top of the other, L's coming just before C's. The coincident production of sound between interactants is interpreted as interpersonal synchrony. C and L act as a single unit here. There is no time for a feedback loop to operate. As the rhythmical transcription in Figure 6.7 indicates, they enter together, on the beat.

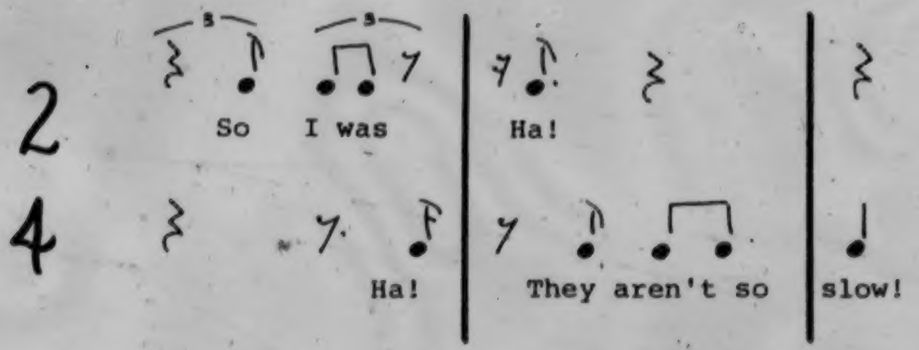
L's burst of laughter--L executes a number of movements coincidentally with her laugh. On the tape the movement looks like a "hiccup": she moves her head up, tilts her head forward, bringing her chin towards her chest, and raises her shoulders. The quick up-down movement anticipates by perhaps 0.1 second, and then overlaps, with a similar up-down motion of the fundamental frequency It is a coinstantaneous complex of sound and movement, interpreted here as intrapersonal synchrony. Action and utterance occur simultaneously, integral parts of the state of play delivered with their occurrence. As before, the very small time frame and precise integration of modalities involved obviates the stimulus-response model.

C's burst of laughter--Although C is in motion through her laugh, movement is largely a continuation, following trajectories of previously initiated motion. The continuation again demonstrates the futility of attempting to establish invariant relationships between movement and utterance. There is one movement sequence of C's left hand that, because it exactly parallels the movement of the fundamental frequency (rising and falling with it), is argued to constitute intrapersonal synchrony.

They aren't so slow--L throws out this timely comment in the middle of C's story. L anticipates a bit, entering the discourse before C can recover from her outburst of laughter, and continue her tale of the "delinquents" at the vocational junior high school. L's verbal contribution redirects the text a bit, but does not interrupt the flow of what is going on. L maintains fundamental frequency, and perpetuates the perceived beat or rhythm of utterance through her utterance. A rhythmic transcription of this portion is provided in Figure 6.8. It is proposed that the fact that L enters the discourse slightly ahead of C and then maintains the established parameters of sound motion through the duration of her utterance is an example of what is here interpreted as that aspect of interpersonal synchrony termed "smooth sequencing of repartee." Coincidentally, such an occurrence also makes a strong argument for a companion notion of interaction as a single system in evolution in which interactants, the creators of the system, immerse themselves, becoming, for the duration of the system, nodes--integral parts of the system in evolution.

L moves with her utterance. Head movements go with the fundamental frequency. The coincident and coetaneous production of sound and motion in one individual is here interpreted as intrapersonal synchrony. There are several instances of note. First L executes a complex of movements coinstantaneously with commencement of her utterance: she shifts her head to the right, right hand to the left, and left hand down and away simultaneously with initiation of "they." Through the balance of her utterance L's head and

Figure 6.8: Musical Notation for "they aren't so slow"



torso seem to "float down" from the height achieved in the "hiccup" of motion accompanying her laugh. The floating goes with and parallels the generally downward motion of the fundamental frequency. The smaller up and down movements of the fundamental frequency embroidering the overall descent are not imitated in head or torso movement. Rather, the small horizontal movements which disturb the descending arc of her head movement correspond to word production. L moves down-left with "aren't," veers down-right with "so," and slips straight down with "slow." The accord between word change and head movement, and the lack of accord between head movement and the smaller shifts of the fundamental frequency demonstrate once more that it is not worthwhile to search for fixed relationships between speech and movement. Sound and motion appear to be variably parcelled out through the sensory modalities. Mair's analogy of a differential gear might be invoked to describe the workings of the "mechanism."

L's hand movements through "they aren't so slow," are similarly coincident. Going with the fundamental frequency, moving down with utterance, they are thought to constitute manifest evidence of intrapersonal synchrony. L's left hand drops quickly with the fundamental frequency of "they," coming to rest palm down on the table top. Her right hand, still holding the cup from which she has yet to sip her tea, moves downward through the entire utterance. The horizontal shifts that punctuate the descending arc of her right hand correspond closely to the smaller changes which embroider the overall descent of the fundamental frequency. Specifically, as the fundamental frequency peaks and begins to fall in "aren't," L's right hand shifts from down-right to down; as the fundamental frequency arches over and down in "so," right hand veers to the left; and with production "slow," L completely reverses the direction of her right hand, making reading for her sip of tea. The entire production takes less than 1.0 second. Speech and movement occur together, a continuous flow of sound and motion in time.

C also moves with L's utterance through this segment. Speech and movement between interactants are coetaneous manifestations in time. C is in the process of completing her loop here. Swinging head and torso clockwise, she moves down, back, and around. The fact that she completes her loop and sets herself up to re-enter the discourse precisely as the fundamental frequency bottoms out and L rounds off her utterance is taken as evidence of interpersonal synchrony. C's head movement, in particular the rotation or turning, is very similar to L's; it too corresponds to and follows word production rather than the smaller up and down variations that embroider the overall descent of the fundamental frequency. C turns her head to the right with "they"; lowers her head, tilting it forward, and looks down, breaking eye contact on "aren't"; moves back to the left and raises her eyes to re-establish eye contact in "so"; and completes her loop, moving toward L and then away as "slow" ends. This is interpreted as an example of interpersonal synchrony in action. C and L move together. And finally, there is the brief gesture of C's right hand. Her hand describes an arc, rising from and descending to the table top. This arc is coinstantaneous with and a perfect imitation of the arc described by the fundamental frequency through "so." Such finely detailed coincidence of movement and sound is argued to support the notion of interpersonal synchrony.

That's right! So I was a real old bitch.

Example VI.3d

The fundamental frequency rises dramatically here and is maintained in a very high frequency range right through the exclamation "that's right." It is, then, an emphatic, high frequency interjection. Mair's model proposes that pitch excursion such as this in fact constitutes a bid for topic control, an effort to bring the shared text, the state of play, back into line with one's own projection of topic. It appears that C is attempting here to override L, to prevent any further interruption, and to finish her story. She is successful. C speaks alone through this and succeeding segments. It is proposed that L responds to the "pitch excursion" in "that's right" and gives over control of the verbal portion of the text, allowing C to complete the story of her experiences at the vocational junior high school. It is important to note that L does not opt out of the interaction here. She continues to move in time with the fundamental frequency of C's utterance throughout; she is still part of the single system in evolution. Apparently she has simply decided to listen for a time.

That's right--C moves emphatically with her utterance. She turns her head abruptly to the right as she produces "that's right," almost throwing out the words. Her hands follow in time, paralleling the direction of movement of the fundamental frequency. C's hands rise from her lap as the fundamental frequency peaks in "right." These simultaneous productions of sound and motion are here interpreted as intrapersonal synchrony.

L also moves with the fundamental frequency, but less dramatically. Her head moves to the right as the fundamental frequency climbs through "that's," and changes direction, veering down, precisely as the fundamental frequency peaks and tumbles over in "right." The coincidence of movement and sound between interactants is taken as evidence of interpersonal synchrony.

So I was--Through this portion of the text C evidences a number of incidences of coincidental and coetaneous production of sound and motion. It is proposed that, individually and collectively, they constitute a strong argument for the existence of intrapersonal synchrony. Specifically, C reverses the direction of her head movement coinstantaneously with the reversal of the direction of motion of the fundamental frequency in "so." Both hands rise with the fundamental frequency. Her right hand peaks with the fundamental frequency, turning with it to move down, toward the table top. Through "I was," C's head and the fundamental frequency contour effect similar kinds of motion: both follow smooth continuous trajectories through time.

L's movements go with the relatively flat intonation contour here as well. That is, they too are relatively "flat," following smooth, even trajectories through time. And it is here, in time with the fundamental frequency of utterance, that L at last brings her cup to her mouth. These instances of coincident and congruent production of sound and motion between interactants are argued to support the notion of interpersonal synchrony.

A real old bitch--The fundamental frequency arches through "a real." C moves with the fundamental frequency. There are two movements of note. First, C changes the direction of her head movement, swerving down, precisely as the fundamental frequency changes direction, turning to fall over and down with "real." Second, the motion of C's right hand is altered coinstantaneously with the alteration of the fundamental in "real." Such simultaneous manifestation of sound and movement within an interactant is interpreted as intrapersonal synchrony.

L moves with the fundamental frequency of "a real" also, even though much of L's movement here has to do with motions (of head and right hand-cup) that are attributable to the process of drinking. Generally these movements are finely detailed, almost too fine for transcription here, and they do not seem to go with the fundamental frequency particularly (or any other notable parameter used to described the interaction). They are not synchronous, then. There is, however, one head movement that is; it quite obviously does go with the movement of the fundamental frequency. L moves her head down with the fundamental frequency in "real." It is the single head movement made along the vertical axis through this portion of the text.

Old bitch--As the fundamental frequency begins to rise through "bitch," C's head moves up and to the right, with it. Her left hand reverses direction coinstantaneously, dropping down-left, off camera. And, in the 0.1 second which marks the end of the utterance, C reaches the apex of the arc described by the movement of her head and, turning her head toward L, re-establishes eye contact. These instances of coincident sound and movement within a single interactant are taken as evidence of intrapersonal synchrony.

There is no interpersonal synchrony in evidence through this portion of the text. L continues to sip at her tea. The transcription indicates small movements of her head and right hand. They do not seem to be especially related to what C is doing or saying. The one notable movement--L moves away from C as C says "bitch"--might be considered coetaneous with the vocal production and thus interpersonal synchrony, but that would seem to require stretching the argument unnecessarily.

And they behaved themselves--The movement of C's head and left hand parallels, going with, the rapidly falling fundamental frequency of "and." This coincidence of sound and motion is thought to demonstrate intrapersonal synchrony. L moves coinstantaneously with the initiation of utterance. Her head and right hand separate, moving in opposite directions: L's head moves up; her right hand drops down. She has finished her sip of tea. The coincidence of speech and movement between interactants is interpreted as interpersonal synchrony. It all happens rapidly and with a finely controlled exactitude that is beyond the capabilities of a feedback, stimulus-response model.

The fact that C drops her left and "clicks" her ring on the table top in time with the established rhythm or pulse of the fundamental frequency of her utterance is thought to be an example of intrapersonal synchrony. It is thought that the concept of interpersonal synchrony receives support from the simultaneous manifestation of sound and motion between interactants here: L moves her hand up with the rising fundamental frequency of "they."

There is very little movement during "behaved," but once more what movement there is goes precisely with the fundamental frequency of utterance. L's head moves down in time with the falling fundamental frequency and as the fundamental bottoms out and begins to rise, both C and L alter the direction of the movement of their heads. C stops horizontal motion--she begins to nod. L ceases vertical motion, shifting to a horizontal (across-left) motion. At the same time L initiates a movement down-left with her right hand. It is a coincidence of sound and motion that is thought to argue strongly for the notion of synchrony in interaction. On C's part the simultaneous manifestation of speech and movement is interpreted as intrapersonal synchrony; on L's part, as interpersonal synchrony.

L moves with the fundamental frequency of "themselves." Head movement precedes to the left and up, going with, paralleling, the fundamental frequency. This congruence of aural and visual modes between interactants is argued to be interpersonal synchrony in action.

And they did some work--As noted above, there is no mingograph record of fundamental frequency for this segment of text. Nevertheless, there is apparent in the transcription of movement one final example of what is here interpreted as interpersonal synchrony. Perhaps fittingly, it ends the fragment. In the same 0.1 second in which C completes her utterance and stops, freezing for an instant, L places her cup (and right hand) down on the table top and pauses, also freezing for an instant. It looks to be one of Mair's "timeless moments." Lasting perhaps 0.1 second, it does not bring the text to a full stop. At best there is a hesitation, a pause. C completes her story, and as she finishes the cognitive model she brings the fundamental frequency down, rendering the physical model stable. This is apparent even without the help of the electrical measuring devices. C brings L with her to this point of stability in time. She makes L as she is, in time. They are there, together, in the realisation of the timeless moment.

L breaks off the moment, redirecting the text into a new topic with her question: "What were you subbing?", and they continue.

Chapter 7

Conclusions

7.1 Introduction

This study began with the observation that there is no coherent paradigm for the study of human interaction. Each tradition and each discipline has its own paradigm, constraining study. Such constraints were claimed to be limiting, in that they prohibit investigation of "what is really going on."

Early studies of human interaction focussed upon language in isolation and were particularly susceptible to such constraints. There were at least two distinct traditions. One addressed language interaction from the perspective of linguistics and the problem of creating context-sensitive grammars for language use. The tradition of generative and interpretive semantics is such a tradition. The other assumed language to be part of human interaction, and focussed upon the rules for language use. Discourse analysis and conversation analysis in anthropology and sociology represent such a tradition. Attempts to synthesize the approaches were made in the work of such scholars as Dell Hymes and John Gumperz.

Urion (1978) has claimed that these attempts failed to capture the salience of human language interaction because they assume the equivalence of "sentence" to "utterance," and because they are constrained, using the model provided in linguistics, to specify "rules" for language production. They are studies of syntax.

This work is part of the tradition that denies that "extralinguistic" factors in human speech production and perception are in fact "extralinguistic." Instead, an ethological approach to human speech interaction demands a focus beyond syntax, rules, and semantics, to a more nearly objective description.

Mair's model, explained in publication and papers during the years from 1977 to 1980, was chosen as the most compelling and comprehensive model with which to examine the process of human speech interaction. Mair proposes a biologically based model to account for the fact of something observed in interaction, something intuitively obvious, and something for which we really lack an appropriate term, but which has aspects of "shared awareness." He attempts to base his model on observable data only, and to eschew "mentalism," "impressionistic interpretation," and retrospective analysis using affective terms. He presumes upon the physiological constraints of the system(s) involved to constrain theorising. His search is for a natural segmentation of language streams; that is, one that is based on perception and immediately verifiable by observation of behaviour. It is what he calls a "natural science" approach to the study of human interaction.

7.2 The Present Study

This study is not a replication of Mair's work in any formal sense. Instead, a central notion, that of synchrony, is examined. The term has a history in the literatures of both linguistics and human ethological studies of interaction, but its definition is controversial. For example, Chapple describes synchrony as the perfect alternation of active-latent phases in interaction, with no overlap of such phases, and with no lacunae in the alternation of the phases. Such a definition of synchrony denies an aspect of synchrony that is integral to Mair's description: the synchrony of sound and movement within a single interactant and a kind of entrainment of one interactant by another, in a realisation of that intrapersonal synchrony.

It is assumed in this study that Mair has provided enough clear criteria for his definition of synchrony that it could be described as an identifiable, observable and definable "event" in human interaction, even if the research protocols did not exactly duplicate Mair's.

The notion of synchrony was chosen because it is central to Mair's model. It does not prove the adequacy of his model, but it is necessary to it: if synchrony is not there, the model fails. Synchrony becomes an ascendant descriptor for this work, a way into the model. The study does not replicate Mair's own study, then, and cannot be said to "prove" the model. At best, it might lend support to it.

This study is a micro-study of three short segments of dyadic human communication, extracted from half-hour speech events. There is an elaborate description for each short segment, detailing hand and head movement and the fundamental frequency trace provided by a mingograph record of the segments. These are described in detail for each interactant, detailing change in each modality and individual at each 0.1 second interval.

There is a mass of data in these 28.2 seconds of interaction. Mere presentation of the data is a major undertaking. However, the sheer bulk of the data is not intended to obfuscate or mystify the study and allow simple assertion of synchrony. Rather, a central concern of the study has been to establish synchrony as a demonstrably observable, definable, and identifiable aspect of human interaction. Microanalysis of the sort employed here might be seen to leave itself open to criticism as overly specific and ungeneralisable, detailing in fact only isolated incidents which may or may not represent events which are characteristic of interaction generally. That is, it might be argued that synchrony as documented in these selected fragments of text belongs only to those bits of text. Both the form and content of the data presentation constitute a response to that criticism. There has been an attempt to document precise integration of audio and visual modes within and between interactants such that the presence of what is here called synchrony may be held to be neither uncommon, unusual, haphazard, nor accidental. Neither is what has been observed explicable by more traditional models of human interaction. They are unable to account for what happens within the time frame of natural spontaneous interaction as it happens.

7.2.1 Contributions of the Study

7.2.1.1 Synchrony

Mair uses the term synchrony. This study has forced a more explicit definition than his. That definition is "the simultaneous manifestation of sound and movement within and between interactants." The definition comes about because of the way in which the study was operationalised differently from Mair's, and because other scholar's definitions of synchrony had to be addressed.

The kinds of synchronies that were observed in the study, potentiated by taking Mair's model as *a priori*, were called interpersonal and intrapersonal synchrony. Intrapersonal synchrony includes coincident movement and sound patterns produced by one individual. Specifically, the study documents simultaneous initiation of sound and movement within a single individual; simultaneous cessation of sound and motion; coetaneous production of sound and movement; coincident shifts in audio and visual modes; parallel movement between modalities; and onomatopoeic movements of audio and visual production. Interpersonal synchrony expands the dimensions of synchrony because it includes coincident patterns of sound and movement between interactants. Interpersonal synchrony documented here involves simultaneous initiation of sound and/or movement between individuals; simultaneous cessation of sound and/or motion between individuals; coetaneous production of sound and/or motion between interactants; production of parallel and/or "mirror image" motion and/or sound between individuals; coincident shifts in audio and/or visual modes between interactants; parallel and/or onomatopoeic motion of audio and/or visual production between individuals; the phenomenon that Mair calls "smooth sequencing of repartee"; and implied simultaneous realisation of cognitive models between interactants.

7.2.1.2 Substantiation of the observation of synchrony

This study substantiates the claim that what is observed and documented must in fact be synchrony and not fortuitous co-occurrence: it happens too often, too regularly--and most important--too precisely and quickly, to be accidentally coincidental. The very small time frame and close integration of modalities within and between interactants obviates claims to the contrary.

7.2.1.3 Single central patterning program in the brain

Mair claims that the speech with movement stream is so closely synchronised that it must be the product of a single central patterning program in the brain. This study cannot address that issue, except in speculative terms. In so far as Mair can make such a claim, it applies to the data of this study as well. At this point, however, one might only conjecture that that accounts for how it is that these "arbitrarily extracted parameters" of communication are manifest so nearly simultaneously, with such close integration of modalities within and between individuals.

7.2.1.4 The social brain

The study, because of the evidence of synchronies, particularly interpersonal synchrony, forces us to look at one system in operation in dyadic communication. It obviates appeal to the stimulus-response model and to the "rule-governed behaviour" models of human interaction. Its implications are in two directions: study of what Mair calls the "social brain," that is, a single system joined by separate sensory modalities in separate individuals; and a study of "social physiology," that is, the physiological and anatomical characteristics of human behaviour. In and of itself documentation of interpersonal synchrony without utterance points to the existence of a supra-individual "beat" running through interaction and argues strongly for the notion of human interaction as a single system in evolution, of which individuals are nodes, organising and organised by it.

7.2.1.5 Autopoiesis

An important implication of the study is in the invocation of the notion of autopoiesis. While the study did not employ the terms of that tradition, and while it was in no sense a "test" of the adequacy of the theory of autopoiesis, a major implication of the study is that the theory put forward by Maturana and Varela, reviewed in Chapter 4, provides categories for description of the processes observed in this study without forcing mechanistic interpretations. It does so because of its requirement to specify and distinguish descriptive domains: interactional, consensual, and linguistic. The dynamic of "structural coupling" fits the observations in this study about synchrony. It would seem that the study itself is about "structural coupling." The notion of structural coupling removes "context" from the morass and easy dodge of calling things "cultural."

7.2.1.6 Incorporation of other models

This study speaks to several other models employed to describe human speech interaction. It incorporates and qualifies Chapple's notion of synchrony. The argument could be extended to obviate appeal to notions such as "entrainment" and to redefine "parallel response" as a more elegantly specified synchrony. Chapple's notion of the pacemaker individual relates to Mair's discussion of control. Both the notions of "control of awareness" and of the "pacemaker individual" become simple assertions in the light of Maturana's and Varela's heuristic description: the data do not show that such notions are useful, and the dynamics involved in autopoiesis make such assertions moral, not scientific, questions.

Scollon, with his notions of "ensemble" and "temporal bonding;" and Erickson, with his specification of interaction rhythms, are incorporated into the larger models. Mair's model, and the dynamics of autopoiesis, contextualize their observations.

Insofar as Lehiste and Ohala are concerned, and their discussions of "chain" and "comb" models of speech production, there would seem to be no imperative to adopt either chain or comb model based on the data presented here. The questions they raise, however, in arguing for either model, are germane to this study. The questions that motivate the argument seem to spring from the same insight that Chapple has about "biological oscillators."

7.2.2 Limitations

The focus in the study was upon people speaking English, albeit Canadian English. Mair has used data about

English speakers. We cannot generalize beyond it, though we have data that show synchrony in two quite different dialects of English.

A major limitation of the study is that the limits of technologies employed in both this study and Mair's do not allow one to describe accurately how interactants get to points of synchrony in conversation. That is surely beyond the scope of this investigation, but it is a major implication for further research. Longer segments of speech will have to be analyzed. The technologies employed also prevent consideration of another important aspect of synchrony, what has elsewhere been described as temporally displaced synchrony. That, too, is surely beyond the scope of this study but remains an implication for further research. More refined measurement and mathematical analysis of text fragments must address such displays of synchrony.

Other technological limitations have to do with the size of the grid employed on the video screen: it may have provided too gross a measure for describing body movement accurately. The reliability and accuracy of the mingograph, particularly as the machine picks up "noise," limited the clear specification of fundamental frequency. The physical set-up, as well, during recording sessions, allowed for limited flexibility of camera angle, and set the size of the interactants on the screen at a proportion that was not easily manipulable.

7.2.3 Implications For Further Research

The most important implication is the need for further study of Mair's model to establish the validity or lack thereof of the other aspects of the model.

It is further suggested that research that employs finer distinctions in time (less than 0.1 second intervals) might provide different insights into synchrony.

There are several notions that demand further investigation. Chapple's "economy of energy in interaction" is intuitively obvious from the video taped data, but there was no way to operationalize the notion in this study. The whole question of "control" as it is used by Mair, and the notion of "pacemaker individuals," which Chapple proposes, requires study.

Finally, this study suggests a need to investigate longer segments of interaction. In the data of the long conversations there is an intuitively obvious "beat," a rise and fall of intensity through each half-hour conversation. This should be formally studied. Everyone who viewed the video records or who dealt with the written transcription saw obvious junctures. There is a major issue to be dealt with, and that is the "natural" segmentation of conversation: what distinctions do we draw as we speak, and what are the larger, "natural" categories.

Bibliography

Ainsworth, W.A.
1976 *Mechanisms of Speech Recognition*. New York: Pergammon Press.

Allen, G.A.
1969 Structure of timing in speech production. Paper read at the Meeting of Acoustical Society of America, San Diego.
1968 The place of rhythm in a theory of language. *Working Papers in Phonetics 10*. Los Angeles: University of California, pp. 60-84.

Argyle, M.
1972 Human social interaction. In R.A. Hinde (editor), *Non-Verbal Communication*. Cambridge, U.K.: Cambridge University Press.

Austin, J.L.
1962 *How To Do Things With Words*. New York: Oxford University Press.

Barlow, J.
1960 Discussion. In P.J. DeCoursey (editor), *Phase Control of Activity in a Rodent*. Cold Spring Harbor Symposia on Quantifiable Biology, Volume 25:54-55. New York: Long Island Biological Association.

Barnwell, T.P.
1971 *An algorithm for segment durations in a reading machine context.* Technical Report No. 479. Cambridge: Research Laboratory of Electronics, Massachusetts Institute of Technology.

Bateson, G.
1972 *Steps to an Ecology of Mind*. New York: Ballantine Books.
1979 *Mind and Nature: A Necessary Unity*. New York: E.P. Dutton.

Bateson, P.P.G. and P.H. Kopfler (editors).
1976 *Perspectives on Ethology*. New York: Plenum Press.

Bauman, R. and J. Sherzer (editors)
1974 *Explorations in the Ethnography of Speaking*. New York: Cambridge University Press.

Beattie, G.W.
1980 The role of language production processes in the organization of behaviour. In B. Butterworth (editor), *Language Production, Volume 1: Speech and Talk*. New York: Academic Press.

Bernstein, N.A.
1967 *The Coordination and Regulation of Movements*. Oxford: Pergamon Press.

Birdwhistell, R.L.
1970 *Kinesics and Context: Essays on Body Motion Communcation*. Philadelphia: University of Pennsylvania Press.

Bloomfield, L.
1933 *Language*. New York: Henry Holt and Co.

Blount, B.G. (editor)
1974 *Language, Culture, and Society*. Cambridge, Massachusetts: Winthrop Publishers.

Bouissac, P.
1976 *Circus and Culture*. Bloomington: Indiana University Press.

Buder, E.
1980 The representation and cognition of rhythm. Unpublished B.A.(Hon.) thesis, Harvard University.

Bullard, J.
1980 A natural science approach to human interaction. Manuscript.

Burling, R.
1970 *Man's Many Voices: Language In Its Cultural Context*. New York: Holt, Rinehart and Winston.

Burridge, K.
1973 *Encountering Aborigines: A Case Study--Anthropology and the Australian Aboriginal*. New York: Pergamon Press.

Butterworth, B.
1980 Evidence from pauses in speech. In B. Butterworth (editor), *Language Production, Volume 1: Speech and Talk*. New York: Academic Press.

Byers, P.
1976 Biological rhythms as information channels in interpersonal communicative behavior. In P.P.G. Bateson and P.H. Kopfler (editors), *Perspectives in Ethology, Volume 2*. New York: Plenum Press.

Campbell, D.T. and J.C. Stanley
1963 *Experimental and Quasi-Experimental Designs For Research*. Chicago: Rand McNally College Publishing.

Carpenter, E. and M. McLuhan
1960 *Explorations in Communication*. Boston: Beacon.

Carroll, J.B. (editor)
1974 *Language, Thought, and Reality: Selected Writings of Benjamin Lee Whorf*. Cambridge: M.I.T. Press.

Chafe, W.
1980 (editor) *The Pear Stories: Cognitive, Cultural, and Linguistic Aspects of Narrative Production*. Norwood, N.J.: Ablex Publishing.
1970 *Meaning and the Structure of Language*. Chicago: University of Chicago Press.

Chance, M.R.A.
1962 An interpretation of some agonistic postures: The role of "cutoff" acts and postures. *Symposium of the Zoological Society of London* 8.

Chapple, E.
1980 The unbounded reaches of anthropology as a research science, and some working hypotheses. *American Anthropologist* 82(4):741-758.
1978 The science of humanics: Multidisciplinary Renaissance of general anthropology. *American Anthropologist* 80(1):42-52.
1949 The interaction chronograph: Its evolution and present application. *Personnel* 25:295-307.
1939 Quantitative analysis of the interaction of individuals. *Proceedings of the National Academy of Science* (U.S.) 25:58-67.

Chomsky, N.
1968 *Language and Mind*. New York: Harcourt, Brace and World.
1965 *Aspects of the Theory of Syntax*. Cambridge: M.I.T. Press.
1957 *Syntactic Structures*. The Hague: Mouton.

Cicourel, A.V.
1973 *Cognitive Sociology: Language and Meaning in Social Interaction*. Harmondsworth: Penguin Books.

Condon, W.S. and W.D. Ogston
1967 A segmentation of behavior. *Journal of Psychiatric Research* 5:221-235.
1966 Sound film analysis of normal and pathological behavior patterns. *Journal of Nervous and Mental Disease* 143(4):338-346.

Crystal, D.
1974 *Linguistics*. New York: Penguin Books.

Darnell, R.
1979 Taciturnity in Native American etiquette: A Cree case. Manuscript.
1978 The social organization of discourse: Some introductory remarks. Introductory paper at Symposium, Annual Meeting, American Anthropological Association.
1974 Correlates of Cree narrative performance. In R. Bauman and J. Sherzer (editors), *Explorations In The Ethnography of Speaking*. New York: Cambridge University Press, pp. 315-336.
1971 The bilingual speech community: A Cree example. In R. Darnell (editor), *Linguistic Diversity in Canadian Society*. Edmonton: Linguistic Research, pp. 155-173.

Darnell, R. and A.L. Vanek
1976 The semantic basis of the animate/inanimate distinction in Cree. *Papers in Linguistics* 9:151-180.

Darwin, C.
1862 *The Expression of Emotions in Man and Animals*. (Reprinted, 1965). Chicago: Chicago University Press.

Denes, P.B. and E.N. Pinson
1963 *The Speech Chain: The Physics and Biology of Spoken Language*. Garden City, New York: Anchor Books-Doubleday.

Douglas, M.
1975 *Implicit Meanings: Essays in Anthropology*. London: Routledge and Kegan Paul.

Eibl-Eibesfeldt, I.
1967 *Ethology: The Biology of Behavior*. New York: Holt, Rinehart and Winston.

Ekman, P. (editor)
1973 *Darwin and Facial Expression*. New York: Academic Pres.

Erickson, F.
1981 *Talking to "The Man": Organization of Communication in School Counselling Interviews*. New York: Academic Press.
1980 *Timing and context in everyday discourse: Implications for the study of referential and social meaning*. Sociolinguistic Working Paper No 67. Austin: Southwest Educational Development Laboratory.

Fant, G. and M.A. Tatham (editors)
1975 Auditory Analysis and Perception of Speech. New York: Academic Press.

Feld, S.
1977 Linguistics and ethnomusicology. Ethnomusicology XVIII(2):197-218.

Fillmore, C.J.
1974 Pragmatics and the description of discourse. In C. Fillmore et al. (editors), BS³: Berkeley Studies in Syntax and Semantics Volume I. Berkeley: Department of Linguistics and Institute of Human Learning, University of California.
1968 The case for case. In E. Bach and R.T. Harms (editors), Universals in Linguistic Theory. New York: Holt, Rinehart and Winston

Fishman, J. (editor)
1968 Readings in the Sociology of Language. The Hague: Mouton.

Fromkin, V. and R. Rodman
1978 An Introduction to Language. New York: Holt, Rinehart and Winston.

Fry, D.B.
1979 The Physics of Speech. Cambridge, U.K.: Cambridge University Press.
1971 Time constants in speech. In L.L. Hammerich, R. Jakobson and E. Zwirner (editors), Form and Substance. Copenhagen: Academisk Forlag.

Gardiner, H.
1972 The Quest For Mind: Piaget, Levi-Strauss, and the Structuralist Movement. New York: Random House.

Givens, D.B.
n.d. Ethological expressions in human expressivity. Manuscript.
1978a Contrasting nonverbal styles in mother-child interaction: Examples from a study of child abuse. Semiotica 24(1,2):33-47.
1978b The nonverbal basis of attraction, flirtation, courtship and seduction. Psychiatry 41(4), November.
1977 Infantile reflexive behaviors and nonverbal communication. Sign Language Studies 16:219-236.

Gleason, H.A., Jr.
1967 An Introduction to Descriptive Linguistics. Revised edition. New York: Holt, Rinehart and Winston.

Goffman, E.
1974 *Frame Analysis*. New York: Harper and Row.
1963 *Stigma: Notes On the Management of Identity*. Englewood Cliffs: Prentice-Hall.

Golani, I.
1976 Homeostatic motor processes in mammalian interactions: A choreography of display. In P.D.G. Bateson and P.H. Kopfer (editors), *Perspectives in Ethology, Volume II*. New York: Plenum Press.

Goldman-Eisler, F.
1980 Psychological mechanisms of speech production as studied through the analysis of simultaneous translation. In B. Butterworth (editor), *Language Production, Volume 1: Speech and Talk*. New York: Academic Press.
1968 *Psycholinguistics: Experiments in Spontaneous Speech*. London: Academic Press.
1958 Speech production and the predictability of words in context. *Quarterly Journal of Experimental Psychology* 10:96-106.

Gregory, R.L.
1970 *The Intelligent Eye*. London: Weidenfeld and Nicolson.

Grimes, J.
1972 The thread of discourse. Manuscript.

Gumperz, J.J. and D. Hymes (editors)
1972 *Directions in Sociolinguistics: The Ethnography of Communication*. New York: Holt, Rinehart and Winston.
1964 The ethnography of communication. *American Anthropologist* 66(6), Part 2.

Gumperz, J.J., T.C. Jupp and C. Roberts
1979 Cross Talk: A Study of Cross-Cultural Communication (A Film and Notes). London: British Broadcasting Corporation and the National Centre for Industrial Language Training.

Gunter, R.
1974 *Sentences in Dialog*. Columbia, South Carolina: Hornbeam Press.

Haas, M.R.
1977 Anthropological linguistics: History. In A.F. Wallace *et al*. (editors), *Perspectives on Anthropology 1976*. American Anthropological Association Special Publication, pp. 33-47.

Hall, E.T.
1959 The Silent Language. Greenwich, Conn.: Fawcett Publications.

Halliday, M.K.
1973 Explorations in the Functions of Language. London: Edward Arnold.
1970 Notes on transitivity and theme in English: II. In J. Lyons (editor), New Horizons in Linguistics. New York: Penguin Books.

Hammerich, L.L., R. Jakobson and E. Zwirner (editors)
1971 Form and Substance. Copenhagen: Academisk Forlag.

Harris, M.
1968 The Rise of Anthropological Theory: A History of Theories of Culture. New York: Thomas Crowell.

Hedley, M.
1972 Acculturation studies of North American Indians: A critique of the underlying framework and its implications. Unpublished M.A. thesis, Department of Anthropology, University of Alberta.

Hess, E.H.
1965 Attitude and pupil size. Scientific American, April.

Hinde, R.A. (editor)
1972 Non-Verbal Communication. Cambridge, U.K.: Cambridge University Press.

Huggins, A.W.F.
1975 On isochrony and syntax. In G. Fant and M.A. Tatham (editors), Auditory Analysis and Perception of Speech. New York: Academic Press.
1972a Just noticeable differences for segment duration in natural speech. Journal of the Acoustic Society of America 51:1270-1278.
1972b On the perception of temporal phenomena in speech. Journal of the Acoustic Society of America 51:1279-1290.

Hymes, D.

1977 The state of the art in linguistic anthropology. In A. Wallace et al. (editors), Perspectives on Anthropology 1976. Special Publication of the American Anthropological Association No. 10.

1974a Ways of speaking. In R. Bauman and J. Sherzer (editors), Explorations in the Ethnography of Speaking. New York: Cambridge University Press, pp. 433-451.

1974b Foundations in Sociolinguistics: An Ethnographic Approach. Philadelphia: University of Pennsylvania Press.

1972 Models of the interaction of language and social life. In J.J. Gumperz and D. Hymes (editors), Directions in Sociolinguistics. New York: Holt, Rinehart and Winston, pp. 35-71.

1972 Reinventing Anthropology. New York: Vintage Books.

1964 (editor) Language in Culture and Personality. New York: Harper and Row.

Jaffe, J. and F. Feldstein

1970 Rhythms of Dialogue. New York: Academic Press.

Jassem, W.

1971 Phonological segmental units in the speech signal. In L.L. Hammerich et al. (editors), Form and Substance. Copenhagen: Academisk Forlag.

Julesz, B.

1971 Foundations of Cyclopean Perception. Chicago: University of Chicago Press.

Keenan, E.O. and B.B. Schieffelin

1976 Topic as discourse: A Study of topic in the conversation of children and adults. In C.N. Li (editor), Subject and Topic. New York: Academic Press, pp. 337-384.

Kendon, A.

1979 Some emerging features of face-to-face interaction studies. Sign Language Studies 22:7-22.

1975 Gesticulation, speech and gesture theory of language origin. Sign Language Studies 2:349-57.

1967 Some functions of gaze-direction in social interaction. Acta Psychologia 26:22-63.

Kendon, A., R.M. Harris and M.R. Kay (editors)

1975 The Organization of Behavior in Face-to-Face Interaction. The Hague: Mouton.

Koerner, E.F.K.

1973 Ferdinand de Saussure. Germany: Friedr. Viewig und Sdm. Gmblt., Braunschweig.

Kozhevnikov, N.A. and L.A. Chistovich

1965 *Speech: Articulation and Perception*. Washington: U.S. Department of Translation, JPRS 30-543.

Kuhn, T.S.

1962 *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press.

Labov, W.

1963 The social motivation of sound change. *Word* 19:273-309.

Labov, W. and D. Fanshell

1977 *Therapeutic Discourse: Psychotherapy as Conversation*. New York: Academic Press.

Lakoff, G.

1974 Pragmatics in natural logic. In C. Fillmore *et al*. (editors), *BS³: Berkeley Studies in Syntax, Volume 1*. Berkeley: Department of Linguistics and Institute of Human Learning, University of California, X-1 to X-46.

Lehiste, I.

1977 Isochrony reconsidered. *Journal of Phonetics* 5:253-263.

1973 Rhythmic units and syntactic units in production and perception. *Journal of the Acoustical Society of America* 54:1228-1234.

1971 Temporal organization of spoken language. In L.L. Hammerich *et al*. (editors), *Form and Substance*. Copenhagen: Academsk Forlag.

1970 *Suprasegmentals*. Cambridge: M.I.T. Press.

Leisman, G. and J. Schwartz

1976 Occular motor variables in reading disorders. In R.M. Knights and D.J. Bakker (editors), *The Neurophysiology of Learning Disorders*. London: University Park Press.

Lenneberg, E.

1967 *Biological Foundations of Language*. New York: Wiley.

1964 (editor) *New Directions in the Study of Language*. Cambridge: M.I.T. Press.

Levi-Strauss, C.

1971 *L'homme nu*. Paris: Librairie Plon.

1966 *The Savage Mind*. Chicago: University of Chicago Press.

1963 *Structural Anthropology*. New York: Basic Books.

Lidov, D.
1972 An example (from Kulintang) of a generative grammar for melody. Paper presented at the Society for Ethnomusicology Meeting, Toronto.

Lieberman, A.M.
1972 Perception of the speech code. In E. David and P. Denes (editors), *Human Communication, A Unified View*. New York: McGraw-Hill.

Lieberman, P.
1967 *Intonation, Perception and Language*. Cambridge: M.I.T. Press.

Lindblom, B. and J. Sundberg
1970 Towards a generative theory of melody. *Swedish Journal of Musicology* 52:71-87.

Linden, E.L.
1974 *Apes, Men, and Language*. New York: Pelican Books.

Lyons, J.
1970a (editor), *New Horizons in Linguistics*. New York: Penguin Books.
1970b *Noam Chomsky*. Glasgow: Fontana/William Collins Sons.
1968 *Introduction to Theoretical Linguistics*. Cambridge: Cambridge University Press.

Maclay, H.
1971 Overview. In D.D. Steinberg *et al*., (editors), *Semantics: An Interdisciplinary Reader In Philosophy, Linguistics, and Psychology*. Cambridge: Cambridge University Press.

Mair, M.W.
1980 A model of the text generator. Paper presented at the International Summer Institute for Structural and Semiotic Studies, Toronto, June.
1978 *Steps Towards Principles of Text Regulation*. Toronto: Toronto Semiotic Circle.
1977 Steps towards principles of text regulation: Papers from a project on speech with movement. Manuscript.

Maturana, H.
1976 The biology of language: The epistemology of reality. In R.W. Rieber (editor), *The Neurophysiology of Language*. New York: Plenum Press.

Maturana, H. and F. Varela
1975 *Autopoetic Systems*. Biological Computer Laboratory Report No. 9.4. Urbana: University of Illinois.

Morris, D.
1977 Manwatching: A Field Guide to Human Behaviour. New York: Abram.

Murray, S.
1979 Social network theory and American sociolinguistics. Unpublished Ph.D. dissertation, University of Toronto.

Murphy, R.F.
1971 The Dialectics of Social Life: Alarms and Excursions in Anthropological Theory. New York: Basic Books.

Nattiez, J.-J.
1974 On the relations between anthropology and the semiotics of music. Paper presented at SEM Meeting, Toronto.

O'Connor, J.D. and G.F. Arnold
1974 Intonation of Colloquial English London: Longmans.

Ohala, J.
1975 The temporal regulation of speech. In G. Fant and M.A. Tatham (editors), Auditory Analysis and Perception of Speech. New York: Academic Press.
1970 Aspects of the control and production of speech. Working Papers in Phonetics 15. Los Angeles: U.C.L.A.

Osgood, C.E.
Dimensionality of the semantic space for communication via facial expression. Scandinavian Journal of Psychology 7.

Pawley, A. and F. Snyder
n.d. Sentence formulation in spontaneous speech: The one-clause-at-a-time hypothesis. Manuscript.

Perkins, L.A.
1980 Toward a modern Pacific aesthetics for poetry. In L.A. Perkins (editor), Kingdoms of the Heart. Honolulu: Kamalu'uluolele.

Philips, S.U.
1974 Warm Springs Indian time: How the regulation of participation affects the progression of events. In R. Bauman and J. Sherzer (editors), Explorations in the Ethnography of Speaking. New York: Cambridge University Press, pp. 92-109.
1972 Participant structures and communicative competence. In C. Cazden et al. (editors), The Functions of Language in the Classroom. New York: Teachers College Press, pp. 370-394.

Pike, K.L.
1945 The Intonation of American English. Michigan: University of Michigan Press.

Powers, H.
1980 Language models and music analysis. Ethnomusicology XVIII(1):1-60.

Ricouer, P.
1973 The model of the text: Meaningful action considered as text. New Literary History 5:91-117.

Ross, J.R.
1975 Where to do things with words. In P. Cole and J. Morgan (editors), Syntax and Semantics 3: Speech Acts. New York: Academic Press.

Sabin, E.J., E.J. Clemmer, D.C. O'Connell and S. Kowal
1979 A pausological approach to speech development. In A. Siegman and S. Feldstein (editors), Of Speech and Time. Hillsdale, N.J.:Lawrence Erlbaum Associates.

Sacks, H.
1974 An analysis of the course of a joke's telling in conversation. In R. Bauman and J. Sherzer (editors), Explorations in the Ethnography of Speaking. New York: Cambridge University Press, pp. 337-353.
1972 On the analyzability of stories by children. In J. Gumperz and D. Hymes (editors), Directions in Sociolinguistics. New York: Holt, Rinehart and Winston, pp. 325-345.

Sacks, H., E.A. Schegloff and G. Jefferson
1972 A simplest systematics for the organization of turn-taking in conversation. Language 50:3-68.

Sadock, J.
1974 Toward a Linguistic Theory of Speech Acts. New York: Academic Press.

Sapir, E.
1974a The unconscious patterning of behavior in society. In B. Blount (editor), Language, Culture and Society. Cambridge: Winthrop, pp. 32-45.
1974b Language. In B. Blount (editor), Language, Culture and Society. Cambridge: Winthrop, pp. 46-66.
1921 Language. New York: Harcourt, Brace and Company.

Schegloff, E.A.
1972 Notes on a conversational practice: Formulating place. In D. Sudnow (editor), Studies in Social Interaction. New York: The Free Press, pp. 75-119.

Schegloff, E.A. and H. Sacks
1974 Opening up closings. In. R. Turner (editor), Ethnomethodology. Baltimore: Penguin.

Scollon, R.
1981a Tempo, density, and silence: Rhythms in ordinary talk. Manuscript reprint from University of Alaska, Fairbanks, Center for Cross Cultural Studies.
1981b The rhythmic integration of ordinary talk. Paper presented at the Georgetown University Roundtable on Languages and Linguistics, March.
1979 (With S. Scollon) Linguistic Convergence: An Ethnography of Speaking at Fort Chipewyan, Alberta. New York: Academic Press.

Searle, J.R.
1975 Indirect speech acts. In P. Cole and J. Morgan (editors), Syntax and Semantics, Volume 3. New York: Academic Press.
1969 Speech Acts: An Essay in the Philosophy of Language. Cambridge: Cambridge University Press.

Seeger, C.
1958 Prescriptive and descriptive music writing. Musical Quarterly 44:184-195.

Sherzer, J. and R. Darnell
1972 Outline guide for the ethnographic description of speech use. In J. Gumperz and D. Hymes (editors), Directions in Sociolinguistics. New York: Holt, Rinehart and Winston, pp. 548-554.

Siegman, A.W. and S. Feldstein (editors)
1979 Of Speech and Time. Hillsdale, N.J.: Lawrence Erlbaum.
1978 Nonverbal Behavior and Communication. Hillsdale, N.J.: Lawrence Erlbaum.

Siegman, A.W., and B. Pope
1965 Effects of question specificity and anxiety-producing messages on verbal fluency in the initial interview. Journal of Personality and Social Psychology 2:522-530.

Sudnow, D.
1979a Ways of the Hand. New York: Bantam Books.
1979b Talk's Body. New York: Alfred A. Knopf.
1972 (editor) Studies in Social Interaction. New York: Free Press.

Thom, R.
1975 Structural Stability and Morphogenesis: An Outline of a General Theory of Models. Reading: Benjamin.

Trevarthen, C.

1977 Descriptive analyses of infant communicative behaviour. In H.R. Schaffer (editor), Studies in Mother-Infant Interaction. London: Academic Press.

1976 Babies dance to adult speech. In R. Lewin (editor), Child Alive. London: Temple Smith.

Uldall, E.T.

Isochronous stresses. In L.L. Hammerich et al. (editors), Form and Substance. Copenhagen: Academisk Forlag.

Umeda, N., and C. H. Coker

1975 Subphonemic variations in American English. In G. Fant and M.A. Tatham (editors), Auditory Analysis and Perception of Speech. New York: Academic Press.

Urion, C.A.

1980 The ethnography of speaking and the description of conversation. Sixth LACUS Forum. Columbia, S.C.: Hornbeam Press.

1978 Control of topic in a bilingual (Cree-English) speech event. Unpublished Ph.D. dissertation, Department of Anthropology, University of Alberta.

1976 Control of topic in bilingual (Cree-English) meetings. Paper read at Living English Conference, California State Polytechnic University, Pomona, California.

1972 Anthropology, linguistics, and structural semantic models. Manuscript.

Vanek, A.L.

1979 A note on context-sensitive grammar. Papers in Linguistics 12(3-4):271-292.

1977 The Czech pronominal system in direct discourse: A special case. Paper read at the American Anthropological Association Meetings, Houston.

Vanek, A.L., and R. Darnell

1979 Semantic structures: Toward a holistic theory of communication. Manuscript.

1978a Semantic structures of human communication. Manuscript.

1978b On the aim and scope of linguistics. Manuscript.

Varela, F.J.

1979 Principles of Biological Autonomy. New York: North Holland.

1976a Not one, not two: Position paper for the mind-body conference. CoEvolution Quarterly 11:62-67.

1976b On observing natural systems. CoEvolution Quarterly 10:26-31.

1975 A calculus for self-reference. International Journal of General Systems 2(5):5-24.

Woodcock, A. and M. Davis
1978 Catastrophe Theory. New York: Avon Books.

Zlatoutsva, R.
1975 Rhythmic structure types in Russian speech. In G. Fant and M.A. Tatham (editors), Auditory Analysis and Perception of Speech. New York: Academic Press, pp. 477-483.